Math 4242 Fall 2016 (Darij Grinberg): midterm 3 pratice problems

Exercise 1. **(a)** Diagonalize the matrix $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$.

(b) Diagonalize the matrix $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$.

(c) Diagonalize the matrix $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$.

(d) Diagonalize the matrix $\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$ for arbitrary n nonzero reals $a_1, a_2, \ldots, a_n$ satisfying $a_1 + a_2 + \cdots + a_n \neq 0$.

[Part **(d)** is supposed to be challenging! Things like this won't be on the exam.]

(e) Can the matrix $\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ be diagonalized? (This is to show that the $a_1 + a_2 + \cdots + a_n \neq 0$ condition in part **(d)** is needed.)

Solution to Exercise 1. In each case, we follow Algorithm 0.1 from problem set #8. I shall go into details in part **(a)**, and then be brief in parts **(b)**, **(c)** and **(e)** (since the method is the same). I will then sketch an approach to part **(d)**.

Let me repeat once again that diagonalization is not a deterministic process (i.e., you have some freedom during the process), so there are several distinct correct answers. Thus, do not be disheartened if my answers don't match yours!

(a) Set $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$.

Step 1: We have $n = 2$ and thus

$$\det(A - xI_n) = \det\begin{pmatrix} 1-x & 2 \\ 1 & 2-x \end{pmatrix} = (1-x)(2-x) - 2 \cdot 1 = x^2 - 3x.$$

Step 2: Now we must find the roots of this polynomial $\det(A - xI_n) = x^2 - 3x$. This is easy: The roots are 0 and 3, since the factorization $\det(A - xI_n) = x^2 - 3x = x(x-3)$ immediately catches the eye. Thus, the eigenvalues of A are 0 and 3. Let me number them $\lambda_1 = 0$ and $\lambda_2 = 3$.

Step 3: Now, we must find a basis of $\operatorname{Ker}(A - \lambda_j I_n)$ for each $j \in \{1, 2\}$. This is a straightforward exercise in Gaussian elimination:

<u>*Computing* $\operatorname{Ker}(A - \lambda_1 I_n)$:</u> We have

$$\operatorname{Ker}\left(A - \underbrace{\lambda_1}_{=0} I_n\right) = \operatorname{Ker}(A - 0I_n) = \operatorname{Ker}\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}.$$

Hence, $\left(\begin{pmatrix} -2 \\ 1 \end{pmatrix}\right)$ is a basis of $\operatorname{Ker}(A - \lambda_1 I_n)$. (This can be easily found by the standard algorithm for finding a basis of a kernel.)

Computing $\operatorname{Ker}(A - \lambda_2 I_n)$: We have

$$\operatorname{Ker}\left(A - \underbrace{\lambda_2}_{=3} I_n\right) = \operatorname{Ker}(A - 3I_n) = \operatorname{Ker}\begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}.$$

Hence, $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$ is a basis of $\operatorname{Ker}(A - \lambda_2 I_n)$. (This can be easily found by the standard algorithm for finding a basis of a kernel.)

Step 4: Now, we concatenate these three bases into one big list $(s_1, s_2, \ldots, s_m)$ of vectors. So this big list is

$$(s_1, s_2) = \left(\underbrace{\begin{pmatrix} -2 \\ 1 \end{pmatrix}}_{\substack{\text{a basis of} \\ \operatorname{Ker}(A - \lambda_1 I_n)}}, \underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\substack{\text{a basis of} \\ \operatorname{Ker}(A - \lambda_2 I_n)}} \right).$$

Thus, $m = 2$, so that $m = n$, and thus A can be diagonalized.

Step 5: Since s_1 belongs to a basis of $\operatorname{Ker}(A - \lambda_1 I_n)$, we have $\mu_1 = \lambda_1 = 0$. Similarly, $\mu_2 = \lambda_2 = 3$.

Step 6: Now, S is the $n \times n$-matrix whose columns are $s_1, s_2, \ldots, s_n$. In other words,

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

Furthermore, Λ is the diagonal matrix whose diagonal entries (from top-left to bottom-right) are $\mu_1, \mu_2, \ldots, \mu_n$. In other words,

$$\Lambda = \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix}.$$

These are the S and Λ we were seeking.

(b) Set $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$.

Step 1: We have $n = 3$ and thus

$$\det(A - xI_n) = \det\begin{pmatrix} 1-x & 2 & 3 \\ 1 & 2-x & 3 \\ 1 & 2 & 3-x \end{pmatrix} = -x^3 + 6x^2$$

(after some computation).

Step 2: Now we must find the roots of this polynomial $\det(A - xI_n) = -x^3 + 6x^2$. This is easy: The roots are 0 and 6, since the factorization $\det(A - xI_n) = -x^3 + 6x^2 = -x^2(x-6)$ immediately catches the eye. Thus, the eigenvalues of A are 0 and 6. Let me number them $\lambda_1 = 0$ and $\lambda_2 = 6$.

Step 3: Now, we must find a basis of $\operatorname{Ker}(A - \lambda_j I_n)$ for each $j \in \{1,2\}$. This is just as straightforward as in part **(a)**, so let me merely give the results:

Computing $\operatorname{Ker}(A - \lambda_1 I_n)$*:* The list $\left(\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right)$ is a basis of $\operatorname{Ker}(A - \lambda_1 I_n)$.

Computing $\operatorname{Ker}(A - \lambda_2 I_n)$*:* The list $\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$ is a basis of $\operatorname{Ker}(A - \lambda_2 I_n)$.

Step 4: Now, we concatenate these three bases into one big list $(s_1, s_2, \ldots, s_m)$ of vectors. So this big list is

$$(s_1, s_2, s_3) = \left(\underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}}_{\substack{\text{a basis of} \\ \operatorname{Ker}(A - \lambda_1 I_n)}}, \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{\substack{\text{a basis of} \\ \operatorname{Ker}(A - \lambda_2 I_n)}} \right).$$

Thus, $m = 3$, so that $m = n$, and thus A can be diagonalized.

Step 5: Since s_1 belongs to a basis of $\operatorname{Ker}(A - \lambda_1 I_n)$, we have $\mu_1 = \lambda_1 = 0$. Since s_2 also belongs to a basis of $\operatorname{Ker}(A - \lambda_1 I_n)$, we have $\mu_2 = \lambda_1 = 0$. Similarly, $\mu_2 = \lambda_2 = 6$.

Step 6: Now, S is the $n \times n$-matrix whose columns are $s_1, s_2, \ldots, s_n$. In other words,

$$S = \begin{pmatrix} -2 & -3 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Furthermore, Λ is the diagonal matrix whose diagonal entries (from top-left to bottom-right) are $\mu_1, \mu_2, \ldots, \mu_n$. In other words,

$$\Lambda = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

These are the S and Λ we were seeking.

(c) The procedure is precisely the same as for parts **(a)** and **(b)**, so let us merely note down the answer:

$$S = \begin{pmatrix} -2 & -3 & -4 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

and

$$\Lambda = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}.$$

(e) Set $A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$.

Step 1: We have $n = 2$ and thus

$$\det(A - xI_n) = \det \begin{pmatrix} 1-x & -1 \\ 1 & -1-x \end{pmatrix} = (1-x)(-1-x) - (-1)\cdot 1 = x^2.$$

Step 2: Now we must find the roots of this polynomial $\det(A - xI_n) = x^2$. This is easy: The only root is 0. Let me number it $\lambda_1 = 0$.

Step 3: Now, we must find a basis of $\operatorname{Ker}(A - \lambda_j I_n)$ for each $j \in \{1\}$. As usual, we can do it by Gaussian elimination, thus obtaining the basis $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$ of $\operatorname{Ker}(A - \lambda_1 I_n)$.

Step 4: Now, we concatenate these three bases into one big list $(s_1, s_2, \ldots, s_m)$ of vectors. This is easy, because there is only one basis. So this big list is

$$(s_1, s_2) = \left(\underbrace{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}_{\substack{\text{a basis of} \\ \operatorname{Ker}(A - \lambda_1 I_n)}} \right).$$

Thus, $m = 1$, so that $m < n$, and thus A **cannot** be diagonalized.

(d) There are various ways to do this. We can try following the same algorithm as before, but the fact that n is variable complicates it significantly – e.g., how do we compute $\det(A - xI_n)$ in full generality? (It is possible, but requires some deeper study of determinants.)

Instead, let me show another way to do so: We guess a general form for S and Λ based on the results we have obtained in parts **(a)**, **(b)** and **(c)**. Namely, let $s = a_1 + a_2 + \cdots + a_n$. We set

$$S = \begin{pmatrix} -a_2 & -a_3 & -a_4 & \cdots & -a_n & 1 \\ a_1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & a_1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & a_1 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_1 & 1 \end{pmatrix}$$

and

$$\Lambda = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & s \end{pmatrix}.$$

Now, we claim that S is invertible and satisfies $A = S\Lambda S^{-1}$. Obviously, once this is proven, it will follow that S and Λ provide a diagonalization of A.

- Let us prove that S is invertible. There are various ways to check that an $n \times n$-matrix is invertible: for example, it suffices to check that its determinant is nonzero, or that its n columns are linearly independent, or that its n rows are linearly independent; or it also suffices to construct an inverse. The one method that seems to work best for our matrix S is by checking that its n rows are linearly independent. So let me do that. (I will then explain why this suffices.)

  Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be n real numbers such that

  $$\lambda_1 \operatorname{row}_1 S + \lambda_2 \operatorname{row}_2 S + \cdots + \lambda_n \operatorname{row}_n S = \overrightarrow{0}. \tag{1}$$

  We shall show that all of $\lambda_1, \lambda_2, \ldots, \lambda_n$ are 0.

  Indeed, based on the way we defined S, we can easily see that

  $$\begin{aligned} &\lambda_1 \operatorname{row}_1 S + \lambda_2 \operatorname{row}_2 S + \cdots + \lambda_n \operatorname{row}_n S \\ &= \left(-\lambda_1 a_2 + \lambda_2 a_1,\ -\lambda_1 a_3 + \lambda_3 a_1,\ \ldots,\ -\lambda_1 a_n + \lambda_n a_1,\ \lambda_1 + \lambda_2 + \cdots + \lambda_n\right). \end{aligned}$$

  Thus,

  $$\begin{aligned} &\left(-\lambda_1 a_2 + \lambda_2 a_1,\ -\lambda_1 a_3 + \lambda_3 a_1,\ \ldots,\ -\lambda_1 a_n + \lambda_n a_1,\ \lambda_1 + \lambda_2 + \cdots + \lambda_n\right) \\ &= \lambda_1 \operatorname{row}_1 S + \lambda_2 \operatorname{row}_2 S + \cdots + \lambda_n \operatorname{row}_n S = \overrightarrow{0} \end{aligned}$$

  (by (1)). In other words, we have

  $$\begin{aligned} -\lambda_1 a_2 + \lambda_2 a_1 &= 0, \\ -\lambda_1 a_3 + \lambda_3 a_1 &= 0, \\ &\cdots, \\ -\lambda_1 a_n + \lambda_n a_1 &= 0, \\ \lambda_1 + \lambda_2 + \cdots + \lambda_n &= 0. \end{aligned}$$

  The first $n-1$ of these equations can be summarized as follows:

  $$-\lambda_1 a_k + \lambda_k a_1 = 0 \qquad \text{for each } k \in \{2, 3, \ldots, n\}.$$

Thus, for each $k \in \{2, 3, \ldots, n\}$, we have

$$\lambda_k a_1 = \lambda_1 a_k. \tag{2}$$

We have just proven the equality (2) for all $k \in \{2, 3, \ldots, n\}$; but it also holds for $k = 1$ (since $\lambda_1 a_1 = \lambda_1 a_1$). Hence, it holds for all $k \in \{1, 2, \ldots, n\}$. In other words, we have

$$\lambda_k a_1 = \lambda_1 a_k \qquad \text{for all } k \in \{1, 2, \ldots, n\}. \tag{3}$$

Now, recall that $\lambda_1 + \lambda_2 + \cdots + \lambda_n = 0$. Multiplying this by a_1, we obtain $(\lambda_1 + \lambda_2 + \cdots + \lambda_n) a_1 = 0$, so that

$$\begin{aligned} 0 &= (\lambda_1 + \lambda_2 + \cdots + \lambda_n) a_1 = \underbrace{\lambda_1 a_1}_{\substack{=\lambda_1 a_1 \\ \text{(by (3))}}} + \underbrace{\lambda_2 a_1}_{\substack{=\lambda_1 a_2 \\ \text{(by (3))}}} + \cdots + \underbrace{\lambda_n a_1}_{\substack{=\lambda_1 a_n \\ \text{(by (3))}}} \\ &= \lambda_1 a_1 + \lambda_1 a_2 + \cdots + \lambda_1 a_n = \lambda_1 (a_1 + a_2 + \cdots + a_n). \end{aligned} \tag{4}$$

Now, recall that $a_1 + a_2 + \cdots + a_n \neq 0$ (by assumption). Hence, we can divide the equality (4) by $a_1 + a_2 + \cdots + a_n$, and thus obtain $0 = \lambda_1$. Hence, $\lambda_1 = 0$.

Now, for each $k \in \{1, 2, \ldots, n\}$, we have

$$\begin{aligned} \lambda_k a_1 &= \underbrace{\lambda_1}_{=0} a_k \qquad \text{(by (3))} \\ &= 0 \end{aligned}$$

and thus $\lambda_k = 0$ (here, we have divided by a_1, which is legitimate because a_1 is nonzero). In other words, all of $\lambda_1, \lambda_2, \ldots, \lambda_n$ are 0.

Now, forget that we fixed $\lambda_1, \lambda_2, \ldots, \lambda_n$. We thus have shown that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are n real numbers such that (1) holds, then all of $\lambda_1, \lambda_2, \ldots, \lambda_n$ are 0. In other words, the n vectors $\operatorname{row}_1 S, \operatorname{row}_2 S, \ldots, \operatorname{row}_n S$ are linearly independent. Hence, they form a basis of their span. Since their span is $\operatorname{Row} S$ (that is, the row space of S), this rewrites as follows: The n vectors $\operatorname{row}_1 S, \operatorname{row}_2 S, \ldots, \operatorname{row}_n S$ form a basis of $\operatorname{Row} S$. Thus, $\dim(\operatorname{Row} S) = n$.

But the rank of a matrix equals the dimension of its row space. Thus, $\operatorname{rank} S = \dim(\operatorname{Row} S) = n$.

Now, in homework set #4, we have seen that an $n \times m$-matrix A is invertible if and only if $\operatorname{rank} A = n = m$. Applying this to n and S instead of m and A, we conclude that the $n \times n$-matrix S is invertible if and only if $\operatorname{rank} S = n = n$. Thus, the $n \times n$-matrix S is invertible (since $\operatorname{rank} S = n = n$).

- Now, it remains to show that $A = S \Lambda S^{-1}$. This is clearly equivalent to $AS = S\Lambda$, so let us show the latter.

We know how the matrices A, S and Λ look like, so we can compute their products easily. We obtain

$$AS = \begin{pmatrix} a_1(-a_2)+a_2a_1 & a_1(-a_3)+a_3a_1 & \cdots & a_1(-a_n)+a_na_1 & a_1+a_2+\cdots+a_n \\ a_1(-a_2)+a_2a_1 & a_1(-a_3)+a_3a_1 & \cdots & a_1(-a_n)+a_na_1 & a_1+a_2+\cdots+a_n \\ a_1(-a_2)+a_2a_1 & a_1(-a_3)+a_3a_1 & \cdots & a_1(-a_n)+a_na_1 & a_1+a_2+\cdots+a_n \\ a_1(-a_2)+a_2a_1 & a_1(-a_3)+a_3a_1 & \cdots & a_1(-a_n)+a_na_1 & a_1+a_2+\cdots+a_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1(-a_2)+a_2a_1 & a_1(-a_3)+a_3a_1 & \cdots & a_1(-a_n)+a_na_1 & a_1+a_2+\cdots+a_n \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & \cdots & 0 & a_1+a_2+\cdots+a_n \\ 0 & 0 & \cdots & 0 & a_1+a_2+\cdots+a_n \\ 0 & 0 & \cdots & 0 & a_1+a_2+\cdots+a_n \\ 0 & 0 & \cdots & 0 & a_1+a_2+\cdots+a_n \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_1+a_2+\cdots+a_n \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 & s \\ 0 & 0 & \cdots & 0 & s \\ 0 & 0 & \cdots & 0 & s \\ 0 & 0 & \cdots & 0 & s \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & s \end{pmatrix}$$

(since $a_1+a_2+\cdots+a_n = s$) and

$$S\Lambda = \begin{pmatrix} 0 & 0 & \cdots & 0 & s \\ 0 & 0 & \cdots & 0 & s \\ 0 & 0 & \cdots & 0 & s \\ 0 & 0 & \cdots & 0 & s \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & s \end{pmatrix}.$$

Comparing these two results clearly yields $AS = S\Lambda$. Multiplying by S^{-1} on the right hand side, we now obtain $A = S\Lambda S^{-1}$.

So our S and Λ form a diagonalization of A.

[*Remark:* In part **(d)** of the exercise, we have assumed that $a_1, a_2, \ldots, a_n$ are nonzero; but in the solution, we have only used the assumption that a_1 is nonzero. Actually, we can avoid making even this assumption; indeed, the matrix A can be diagonalized even when a_1 is zero, although **the matrices S and Λ would have to be chosen differently in this situation**. On the other hand, the assumption that $a_1+a_2+\cdots+a_n$ is nonzero is crucial; if $a_1+a_2+\cdots+a_n = 0$, then A cannot be diagonalized (unless $a_1 = a_2 = \cdots = a_n = 0$).] □

Exercise 2. A 3×3-matrix $\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ is said to be *magic* if it satisfies the chain of equalities

$$\begin{aligned} &a_1+b_1+c_1 = a_2+b_2+c_2 = a_3+b_3+c_3 \\ &= a_1+a_2+a_3 = b_1+b_2+b_3 = c_1+c_2+c_3 \\ &= a_1+b_2+c_3 = c_1+b_2+a_3. \end{aligned}$$

(When the entries of the matrix are integers, it is what is commonly called a "magic square".) For example, $\begin{pmatrix} -3 & 8 & -2 \\ 2 & 1 & 0 \\ 4 & -6 & 5 \end{pmatrix}$ is a magic 3×3-matrix.

The magic 3×3-matrices form a subspace of $\mathbb{R}^{3\times 3}$. Denote this subspace by $\mathcal{M}_3$.

(a) Find a basis of this space.

(b) Consider the map $\operatorname{col}_1 : \mathcal{M}_3 \to \mathbb{R}^3$ that sends any magic 3×3-matrix to its first column. (Thus, $\operatorname{col}_1 \left(\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \right) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$.)

Is col_1 linear? If so, find the matrix representing it with respect to your chosen basis of $\mathcal{M}_3$ and the standard basis (e_1, e_2, e_3) of $\mathbb{R}^3$.

Is col_1 injective? Surjective? Bijective?

(**Note:** The question whether col_1 is injective is tantamount to asking whether a magic 3×3-matrix is uniquely determined by its first column. For example, if you know that the first column of a magic 3×3-matrix is $\begin{pmatrix} 3 \\ 1 \\ 7 \end{pmatrix}$, then can you reconstruct the whole matrix?

The question whether col_1 is surjective is tantamount to asking whether every column vector of size 3 appears as a first column of a magic 3×3-matrix. For example, is there a magic 3×3-matrix with first column $\begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}$?)

(c) Consider the map $\operatorname{NW}_2 : \mathcal{M}_3 \to \mathbb{R}^{2\times 2}$ that sends any magic 3×3-matrix to its "northwestern 2×2-submatrix":

$$\operatorname{NW}_2 \left(\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \right) = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$

Is NW_2 linear? If so, find the matrix representing it with respect to your chosen basis of $\mathcal{M}_3$ and the standard basis $(E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ of $\mathbb{R}^{2\times 2}$.

Is NW_2 injective? Surjective? Bijective?

(d) Consider the map $\operatorname{mid} : \mathcal{M}_3 \to \mathbb{R}$ that sends any magic 3×3-matrix to its "middle entry":

$$\operatorname{mid} \left(\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \right) = b_2.$$

Is mid linear? If so, find the matrix representing it with respect to your chosen basis of $\mathcal{M}_3$ and the standard basis (1) of $\mathbb{R}$.

Is mid injective? Surjective? Bijective?

(e) The *Siamese map* $S : \mathbb{R}^2 \to \mathcal{M}_3$ is defined as follows:

$$S\left((a,b)^T\right) = \begin{pmatrix} a+7b & a & a+5b \\ a+2b & a+4b & a+6b \\ a+3b & a+8b & a+b \end{pmatrix}.$$

(This corresponds to the Siamese method of constructing magic squares.)

Is S linear? If so, find the matrix representing it with respect to the standard basis (e_1, e_2) of $\mathbb{R}^2$ and your chosen basis of $\mathcal{M}_3$.

(f) The map $T : \mathcal{M}_3 \to \mathcal{M}_3$ is defined as follows:

$$T(A) = A^T.$$

(In other words, it sends a magic 3×3-matrix to its transpose. Notice that the transpose is magic, too, as you can easily see.)

Is T linear? If so, find the matrix representing it with respect to your chosen basis of $\mathcal{M}_3$ and your chosen basis of $\mathcal{M}_3$.

Solution to Exercise 2 (sketched). **(a)** The elements of $\mathcal{M}_3$ are the magic 3×3-matrices, i.e., the 3×3-matrices $\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ satisfying the system of linear equations

$$\begin{cases} a_1 + b_1 + c_1 = a_2 + b_2 + c_2; \\ a_2 + b_2 + c_2 = a_3 + b_3 + c_3; \\ a_3 + b_3 + c_3 = a_1 + a_2 + a_3; \\ a_1 + a_2 + a_3 = b_1 + b_2 + b_3; \\ b_1 + b_2 + b_3 = c_1 + c_2 + c_3; \\ c_1 + c_2 + c_3 = a_1 + b_2 + c_3; \\ a_1 + b_2 + c_3 = c_1 + b_2 + a_3 \end{cases} . \tag{5}$$

This system (5) can be solved by Gaussian elimination (like any other system). There are various ways the answer can look like, depending on the choices made in the elimination process (i.e., which row operations to perform, and how to order the variables and equations). I am going to "cheat" and present what is perhaps the simplest way to parametrize the solutions:

$$\begin{aligned} &(a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3)^T \\ &= (c-b, c+a+b, c-a, c-a+b, c, c+a-b, c+a, c-a-b, c+b)^T \end{aligned}$$

for $a, b, c \in \mathbb{R}$ arbitrary. In other words,

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} c-b & c+a+b & c-a \\ c-a+b & c & c+a-b \\ c+a & c-a-b & c+b \end{pmatrix} \tag{6}$$

for $a, b, c \in \mathbb{R}$ arbitrary. This is Lucas's method for constructing a magic square of order 3. (Lucas went further and found the precise conditions on a, b, c to ensure that the entries of the matrix are pairwise distinct positive integers, which is what one usually means when speaking of a "magic square".)

We can rewrite (6) as follows:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = \begin{pmatrix} c-b & c+a+b & c-a \\ c-a+b & c & c+a-b \\ c+a & c-a-b & c+b \end{pmatrix}$$
$$= aP + bQ + cR,$$

where

$$P = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}, \qquad Q = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix},$$
$$R = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Thus,

$$\mathcal{M}_3 = \operatorname{span}(P, Q, R).$$

Moreover, it is easy to see that the matrices P, Q, R are linearly independent[1]. Thus, (P, Q, R) is a basis of $\mathcal{M}_3$. We shall denote this basis by $\mathbf{v}$.

Of course, you may have obtained a different basis. (Needless to say, it should have size 3.)

(b) The map col_1 is linear, and the matrix representing it is

$$M_{\mathbf{v},\mathbf{e},\operatorname{col}_1} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

(where $\mathbf{e}$ denotes the standard basis (e_1, e_2, e_3) of $\mathbb{R}^3$).

[Note that the columns of $M_{\mathbf{v},\mathbf{e},\operatorname{col}_1}$ are precisely the first columns of P, Q, R. This is because the coordinates of a vector in $\mathbb{R}^3$ with respect to the basis $\mathbf{e}$ are simply the entries of this vector.]

The map col_1 is bijective, since $\operatorname{rank}\left(M_{\mathbf{v},\mathbf{e},\operatorname{col}_1}\right) = 3 = 3$. Thus, col_1 is injective and surjective.

[1] In fact, if $a, b, c \in \mathbb{R}$ are such that $aP + bQ + cR = 0$, then

$$0 = aP + bQ + cR = \begin{pmatrix} c-b & c+a+b & c-a \\ c-a+b & c & c+a-b \\ c+a & c-a-b & c+b \end{pmatrix},$$

from which it follows that the entries $c - b$, c and $c + a$ are all 0, from which it easily follows that $a = b = c = 0$.

(c) The map NW_2 is linear, and the matrix representing it is

$$M_{\mathbf{v},\mathbf{f},\operatorname{NW}_2} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

(where $\mathbf{f}$ denotes the standard basis $(E_{1,1}, E_{1,2}, E_{2,1}, E_{2,2})$ of $\mathbb{R}^{2\times 2}$).

The map NW_2 is injective, since $\operatorname{rank}\left(M_{\mathbf{v},\mathbf{f},\operatorname{NW}_2}\right) = 3$. But it is not surjective, since $\operatorname{rank}\left(M_{\mathbf{v},\mathbf{f},\operatorname{NW}_2}\right) \neq 4$. Hence, it is not bijective either.

(d) The map mid is linear, and the matrix representing it is

$$M_{\mathbf{v},\mathbf{g},\operatorname{mid}} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$$

(where $\mathbf{g}$ denotes the standard basis (1) of $\mathbb{R}$).

The map mid is surjective, since $\operatorname{rank}\left(M_{\mathbf{v},\mathbf{g},\operatorname{mid}}\right) = 1$. But it is not injective, since $\operatorname{rank}\left(M_{\mathbf{v},\mathbf{g},\operatorname{mid}}\right) \neq 3$. Hence, it is not bijective either.

(e) The map S is linear, and the matrix representing it is

$$M_{\mathbf{h},\mathbf{v},S} = \begin{pmatrix} 0 & -1 \\ 0 & -3 \\ 1 & 4 \end{pmatrix}$$

(where $\mathbf{h}$ denotes the standard basis (e_1, e_2) of $\mathbb{R}^2$).

The map S is injective, since $\operatorname{rank}\left(M_{\mathbf{h},\mathbf{v},S}\right) = 2$. But it is not surjective, since $\operatorname{rank}\left(M_{\mathbf{h},\mathbf{v},S}\right) \neq 3$. Hence, it is not bijective either.

(f) The map T is linear, and the matrix representing it is

$$M_{\mathbf{v},\mathbf{v},T} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The map T is bijective, since $\operatorname{rank}\left(M_{\mathbf{v},\mathbf{v},T}\right) = 3 = 3$. Thus, T is injective and surjective. (Of course, the bijectivity of T can be seen in a simpler way as well: T is invertible, its inverse being itself. After all, every matrix A satisfies $\left(A^T\right)^T = A$.) □

Eigenvectors and eigenvalues make sense not only for square matrices, but also for linear maps from a vector space to itself. (This is not that surprising, seeing that square matrices are used to represent such linear maps.) Here is how they are defined:

Definition 0.1. Let V be a vector space, and $F : V \to V$ a linear map. (We work with real numbers in this exercise, so all vector spaces are over $\mathbb{R}$, and all scalars are in $\mathbb{R}$. The downside of this is that we miss some eigenvalues; but that's OK for an introduction.)

- Given a scalar $\lambda \in \mathbb{R}$ and a vector $v \in V$, we say that v is an *eigenvector of F for eigenvalue* λ if and only if we have $F(v) = \lambda v$.
- A scalar $\lambda \in \mathbb{R}$ is said to be an *eigenvalue* of F if and only if there exists a nonzero eigenvector of F for eigenvalue λ. (Of course, the zero vector $\overrightarrow{0}$ is an eigenvector of F for eigenvalue λ for any $\lambda \in \mathbb{R}$; thus, in order to have an interesting notion, we must require "nonzero".)

One way to find eigenvalues and eigenvectors of linear maps is the following: Let $\mathbf{v}$ be any basis of V (assuming that V is finite-dimensional). Then, the eigenvalues of F are precisely the eigenvalues of the matrix $M_{\mathbf{v},\mathbf{v},F}$. Moreover, the eigenvectors of F for a given eigenvalue λ are the images of the eigenvectors of $M_{\mathbf{v},\mathbf{v},F}$ for λ under the map $L_{\mathbf{v}}$. (See homework set #7 for the definition of $L_{\mathbf{v}}$.)

Example 0.2. Let P_2 be the vector space of all polynomials of degree ≤ 2. Let $F : P_2 \to P_2$ be the map that sends every polynomial $f \in P_2$ to $\left(f\left[x^2\right]\right)''$ (that is, the second derivative of $f\left[x^2\right]$). Here, I am again using the notation $f[y]$ for "f evaluated at y" (which would usually be denoted by $f(y)$, but that would risk being mistaken for a product).

The map F is well-defined (this is easy to check: evaluating a polynomial f at x^2 raises the degree of f to 4 (in the worst case), but then differentiating it twice pulls it back to ≤ 2) and linear (this is, again, straightforward to check). For example,

$$F\left(x^2\right) = \left(\underbrace{x^2\left[x^2\right]}_{=\left(x^2\right)^2=x^4}\right)'' = \left(x^4\right)'' = \left(4x^3\right)' = 12x^2. \qquad (7)$$

What are the eigenvalues and eigenvectors of F ?

Actually, we have seen one of these already: The equality (7) shows that x^2 is an eigenvector of F for eigenvalue 12. But there are more; let's find them all.

We let $\mathbf{v}$ be the basis $\left(1, x, x^2\right)$ of P_2. Then, it is straightforward to compute

$$M_{\mathbf{v},\mathbf{v},F} = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 12 \end{pmatrix}.$$

Now, what are the eigenvalues and the eigenvectors of this matrix $M_{\mathbf{v},\mathbf{v},F}$?

It is easy to see that $\det\left(M_{\mathbf{v},\mathbf{v},F} - xI_3\right) = x^2(12 - x)$. Thus, the eigenvalues of $M_{\mathbf{v},\mathbf{v},F}$ are 0 and 12. Therefore, these are also the eigenvalues of F. Moreover:

- The eigenvectors of $M_{\mathbf{v},\mathbf{v},F}$ for eigenvalue 0 are the elements of

$$\operatorname{Ker}\left(M_{\mathbf{v},\mathbf{v},F} - 0I_3\right) = \operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right),$$

i.e., the multiples of $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

- The eigenvectors of $M_{\mathbf{v},\mathbf{v},F}$ for eigenvalue 12 are the elements of

$$\operatorname{Ker}(M_{\mathbf{v},\mathbf{v},F} - 12I_3) = \operatorname{span}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right),$$

i.e., the multiples of $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

(This, by the way, shows that $M_{\mathbf{v},\mathbf{v},F}$ is not diagonalizable – since it does not have 3 linearly independent eigenvectors.)

Now, recall that the eigenvectors of F for a given eigenvalue λ are the images of the eigenvectors of $M_{\mathbf{v},\mathbf{v},F}$ for λ under the map $L_{\mathbf{v}}$. Thus:

- The eigenvectors of F for eigenvalue 0 are the elements of

$$L_{\mathbf{v}}\left(\operatorname{span}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)\right),$$

i.e., the multiples of

$$L_{\mathbf{v}}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 = 1.$$

In other words, they are the constant polynomials.

- The eigenvectors of F for eigenvalue 12 are the elements of

$$L_{\mathbf{v}}\left(\operatorname{span}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)\right),$$

i.e., the multiples of

$$L_{\mathbf{v}}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = 0 \cdot 1 + 0 \cdot x + 1 \cdot x^2 = x^2.$$

Exercise 3. **(a)** Find the eigenvalues and the eigenvectors of the map $T_e : P_2 \to P_2$ defined by $T_e(f) = x^2 f\left[\frac{1}{x}\right]$. (This map was discussed in Exercise 2 **(e)** of homework set #6.)

(b) Find the eigenvalues and the eigenvectors of the map $T : \mathcal{M}_3 \to \mathcal{M}_3$ from Exercise 2 **(f)**.

(c) Consider the map $T' : \mathbb{R}^{3\times 3} \to \mathbb{R}^{3\times 3}$ that sends each matrix $A \in \mathbb{R}^{3\times 3}$ to its transpose A^T. Find the eigenvalues and the eigenvectors of this map.

[It might be good not to approach part **(c)** this exercise using the $M_{\mathbf{v},\mathbf{v},F}$ method, but rather think in terms of the definition. An eigenvector of the map T' would be a nonzero matrix A whose transpose is a scalar multiple of A. How would such a matrix looks like? What relations would have to hold between its entries?]

Solution to Exercise 3. **(a)** We follow the method from Example 0.2:

We let $\mathbf{v}$ be the basis $\left(1, x, x^2\right)$ of P_2. Then, it is straightforward to compute

$$M_{\mathbf{v},\mathbf{v},T_e} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Now, what are the eigenvalues and the eigenvectors of this matrix $M_{\mathbf{v},\mathbf{v},T_e}$?

It is easy to see that $\det\left(M_{\mathbf{v},\mathbf{v},T_e} - xI_3\right) = -\left(x-1\right)^2\left(x+1\right)$. Thus, the eigenvalues of $M_{\mathbf{v},\mathbf{v},T_e}$ are 1 and -1. Therefore, these are also the eigenvalues of T_e. Moreover:

- The eigenvectors of $M_{\mathbf{v},\mathbf{v},T_e}$ for eigenvalue 1 are the elements of

$$\operatorname{Ker}\left(M_{\mathbf{v},\mathbf{v},T_e} - 1I_3\right) = \operatorname{Ker}\begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix} = \operatorname{span}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right),$$

  i.e., the linear combinations of $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

- The eigenvectors of $M_{\mathbf{v},\mathbf{v},T_e}$ for eigenvalue -1 are the elements of

$$\operatorname{Ker}\left(M_{\mathbf{v},\mathbf{v},T_e} - (-1)\,I_3\right) = \operatorname{Ker}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \operatorname{span}\left(\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right),$$

  i.e., the multiples of $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

Now, recall that the eigenvectors of T_e for a given eigenvalue λ are the images of the eigenvectors of $M_{\mathbf{v},\mathbf{v},T_e}$ for λ under the map $L_{\mathbf{v}}$. Thus:

- The eigenvectors of T_e for eigenvalue 0 are the elements of

$$L_{\mathbf{v}}\left(\operatorname{span}\left(\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}1\\0\\1\end{pmatrix}\right)\right),$$

  i.e., the linear combinations of

$$L_{\mathbf{v}}\left(\begin{pmatrix}0\\1\\0\end{pmatrix}\right)=0\cdot 1+1\cdot x+0\cdot x^2=x$$

  and

$$L_{\mathbf{v}}\left(\begin{pmatrix}1\\0\\1\end{pmatrix}\right)=1\cdot 1+0\cdot x+1\cdot x^2=1+x^2.$$

  In other words, they are the polynomials of the form $b+ax+bx^2$ for $a,b\in\mathbb{R}$.

- The eigenvectors of T_e for eigenvalue -1 are the elements of

$$L_{\mathbf{v}}\left(\operatorname{span}\left(\begin{pmatrix}-1\\0\\1\end{pmatrix}\right)\right),$$

  i.e., the multiples of

$$L_{\mathbf{v}}\left(\begin{pmatrix}-1\\0\\1\end{pmatrix}\right)=(-1)\cdot 1+0\cdot x+1\cdot x^2=-1+x^2.$$

  In other words, they are the polynomials of the form $-a+ax^2$ for $a\in\mathbb{R}$.

(b) Recall the basis $\mathbf{v}=(P,Q,R)$ we chose in the solution of Exercise 2. In the solution to Exercise 2 **(f)**, we found that

$$M_{\mathbf{v},\mathbf{v},T}=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

From this, it is easy to see (as in part **(a)**) that the eigenvalues of T are 1 and -1, and furthermore the eigenvectors are as follows:

- The eigenvectors of T for eigenvalue 1 are the linear combinations of Q and R. In other words, they are the magic 3×3-matrices of the form

$$\begin{pmatrix}b-a&a+b&b\\a+b&b&b-a\\b&b-a&a+b\end{pmatrix}\qquad\text{with } a,b\in\mathbb{R}.$$

- The eigenvectors of T for eigenvalue 1 are the linear combinations of P. In other words, they are the magic 3×3-matrices of the form

$$\begin{pmatrix} 0 & a & -a \\ -a & 0 & a \\ a & -a & 0 \end{pmatrix} \qquad \text{with } a \in \mathbb{R}.$$

Actually, the whole process here is easier than for part **(a)**, because the matrix $M_{\mathbf{v},\mathbf{v},T}$ is diagonal. Finding the eigenvalues and the eigenvectors of a diagonal matrix is particularly easy:

Proposition 0.3. Let A be a diagonal $n \times n$-matrix. Let $a_1, a_2, \ldots, a_n$ be the diagonal entries of A (more precisely, let a_k be the (k,k)-th entry of A). Then:

- The eigenvalues of A are $a_1, a_2, \ldots, a_n$.
- If $\lambda \in \mathbb{C}$, then the eigenvectors of A for eigenvalue λ are precisely the linear combinations of the standard basis vectors e_k for those $k \in \{1, 2, \ldots, n\}$ satisfying $a_k = \lambda$.

We leave the proof to the reader.
(c) Answer:

- The eigenvalues are 1 and -1.
- The eigenvectors for eigenvalue 1 are the symmetric 3×3-matrices (i.e., the matrices of the form $\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$ with $a, b, c, d, e, f \in \mathbb{R}$).
- The eigenvectors for eigenvalue -1 are the skew-symmetric 3×3-matrices (i.e., the matrices of the form $\begin{pmatrix} 0 & b & c \\ -b & 0 & e \\ -c & -e & 0 \end{pmatrix}$ with $b, c, e \in \mathbb{R}$).

Proof. It is best to proceed directly using the definition of eigenvectors:
First, let us see what eigenvalues are possible. Indeed, let λ be an eigenvalue of T'. Then, there exists a nonzero eigenvector $A \in \mathbb{R}^{3\times 3}$ of T' for eigenvalue λ. (The reason why I am calling it A rather than v is that it is a 3×3-matrix, and we're just more used to denoting matrices by capital letters.)
The definition of T' yields $T'(A) = A^T$. But $T'(A) = \lambda A$ (since A is an eigenvector of T' for eigenvalue λ). Thus, $A^T = T'(A) = \lambda A$. Hence,

$$\left(A^T\right)^T = (\lambda A)^T = \lambda \underbrace{A^T}_{=\lambda A} = \lambda\lambda A = \lambda^2 A.$$

Since $\left(A^T\right)^T = A$, this rewrites as $A = \lambda^2 A$. Since A is nonzero, this yields $\lambda^2 = 1$ (why?). Therefore, λ is either 1 or -1.

Thus, we have shown that each eigenvalue of T' is either 1 or -1. It is easy to see that both 1 and -1 actually **are** eigenvalues. (Indeed, we only need to check that at least one nonzero eigenvector exists for each of them. This is easy: $E_{1,1}$ is a nonzero eigenvector of T' for eigenvalue 1, whereas $E_{1,2} - E_{2,1}$ is a nonzero eigenvector of T' for eigenvalue -1.)

It remains to compute the eigenvectors. But this is, again, easy: The eigenvectors of T' for eigenvalue -1 are the matrices $A \in \mathbb{R}^{3\times 3}$ satisfying $T'(A) = (-1)A$. In other words, they are the matrices $A \in \mathbb{R}^{3\times 3}$ satisfying $A^T = -A$ (because $T'(A) = A^T$ and $(-1)A = -A$). In other words, they are the skew-symmetric 3×3-matrices. Similarly, the eigenvectors of T' for eigenvalue 1 are the symmetric 3×3-matrices. $\square$