Comments

The purpose of this practice exercise set is to give a review of the midterm material via easy proof questions. You can read it as a summary and you should attempt to prove each question on your own or with your study group. Notice that one of the most important theorems which we covered in the previous weeks is that if the dimension of a vector space is n, then it is isomorphic to $\mathbb{R}^n$. We saw that this allows us to carry all our computations to the world of MAT223 where everything could be computed by a computer. In the first term test, you didn't have any computational questions because you needed to understand in which spaces you can do computations. In the second term test, you should naturally expect to do some computations. However, you should understand the theoretical background in order to apply the computation techniques from MAT223 to our course. I did not break this into different sections because I wanted to point out that everything is connected. That's why it might seem long and scary, but I believe you could finish this review in 3-4 hours with a good reading.

Linear Transformations Between Finite Dimensional Vector Spaces

Throughout this problem set V and W are finite dimensional vector spaces and $f : V \to W$ is a linear function. We pick a basis $\alpha = \{v_1, \ldots, v_n\}$ for V and a basis $\{w_1, \ldots, w_m\}$ for W.

1. Write what it means for f to be linear.

2. Here a couple of easy consequences of the definition of linear functions:

   (a) Let 0_V and 0_W denote the zero vectors of V and W respectively. Then,

   $$f(0_V) = 0_W.$$

   Note. We don't actually need the subscripts and we will not use them from now on. We will simply write $f(0) = 0$. From the context, you can understand the first zero is the zero vector in V and the second one is the zero vector in W. Indeed, the inputs of f are from V and the inputs of f are from W.
   Hint. To prove this question, start with $0 + 0 = 0$, apply f to both sides and use linearity.

   (b) For any $v \in V$,

   $$f(-v) = -f(v).$$

3. Show that for any $x_1, \ldots, x_k \in V$ and $x_1, \ldots, x_k \in \mathbb{R}$, we have

   $$f(c_1x_1 + \ldots + c_kx_k) = c_1f(x_1) + \ldots + c_kf(x_k).$$

   Recall that since α is a basis, we can write any vector v in V as a linear combination

   $$v = c_1v_1 + \ldots + c_nv_n.$$

   So, you proved that

   $$f(v) = f(c_1v_1 + \ldots + c_nv_n) = c_1f(v_1) + \ldots + c_nf(v_n).$$

   This means that as long as you know $f(v_1), \ldots, f(v_n)$, you can compute $f(v)$ for any v in V. This is one of the miracles of linear functions.

4. Recall from MAT223 that if A is the matrix

$$A = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{mn} \end{bmatrix}$$

and if e_1 is the vector

$$\begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

then Ae_1 is the first column of A. If you don't recall this, convince yourself by computing

$$\begin{bmatrix} 5 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

You should get

$$\begin{bmatrix} 5 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Similarly, you get

$$\begin{bmatrix} 5 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Check that these computations are correct.

5. We will mimick what we did in the previous question to get a matrix which represents f. The recipe is as follows: Compute $f(v_1), \dots, f(v_n)$. Since these are elements of W, they can be written as a linear combination of $w_1, \dots, w_m$:

$$\begin{aligned} f(v_1) &= a_{11}w_1 + \ldots + a_{m1}w_m \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ f(v_n) &= a_{1n}w_1 + \ldots + a_{mn}w_m \end{aligned}$$

looking back at the previous question, we define the matrix of f with respect to α and β as

$$[f]_\alpha^\beta = A = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{mn} \end{bmatrix}$$

Convince yourself that if we choose different bases, the matrix might change. So, this matrix depends on the basis you choose. This is important.

6. The previous question allowed us to store the entire information about the function f using an $n \times m$ matrix. Now, we will store the information about our vectors using vectors. To this end, we will define the coordinate vectors. For $v = c_1v_1 + \ldots + c_nv_n$ in V, we will write

$$[v]_\alpha = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Similarly, for $w = d_1w_1 + \ldots + d_mw_m$ in W, we put

$$[w]_\beta = \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix}$$

We have the following very important consequence of this symbolic definition.

$$[f]_\alpha^\beta [v]_\alpha = [f(v)]_\beta.$$

This can be seen by computing

$$\begin{aligned} f(v) &= f(c_1v_1 + \ldots + c_nv_n) \\ &= c_1f(v_1) + \ldots + cf(v_n) \\ &= c_1[a_{11}w_1 + \ldots + a_{m1}w_m] + \ldots + c_n[a_{1n}w_1 + \ldots + a_{mn}w_m] \\ &= (c_1a_{11} + \ldots + c_na_{1n})w_1 + \ldots + (c_1a_{m1} + \ldots + c_na_{mn})w_m \end{aligned}$$

So,

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \mapsto \begin{bmatrix} c_1a_{11} + \ldots + c_na_{1n} \\ \vdots \\ c_1a_{m1} + \ldots + c_na_{mn} \end{bmatrix}$$

You should see this on a small example where $n = 2$ and $m = 3$, create your own example. This is the key observation, so make sure you understand how this works. You should do lots of computations of this sort, you should get your hands dirty.

7. Most likely you didn't do your own small example which was suggested at the end of the previous question. So, here is a small example. Let $V = P_1(\mathbb{R})$ and $W = \mathbb{R}^3$. Define $f : V \to W$ by

$$f(a + bx) = (a + b, a - b, 2a + 3b)$$

Let $\alpha = \{1, 1 + x\}$ and $\beta = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$. Then,

$$\begin{aligned} f(1) = (1, 1, 2) &= 2 \cdot (1, 1, 1) + (-1) \cdot (1, 1, 0) + 0 \cdot (1, 0, 0) \\ f(1 + x) = (2, 0, 5) &= 5 \cdot (1, 1, 1) + (-5) \cdot (1, 1, 0) + 2 \cdot (1, 0, 0). \end{aligned}$$

Hence,

$$[f]_\alpha^\beta = \begin{bmatrix} 2 & 5 \\ -1 & -5 \\ 0 & 2 \end{bmatrix}$$

Now, consider $3 + 5x$.

(a) What is $[3 + 5x]_\alpha$?

(b) What is $f(3 + 5x)$?

(c) What is $[f(3 + 5x)]_\beta$?

(d) Is it true that

$$[f(3 + 5x)]_\beta = [f]_\alpha^\beta [3 + 5x]_\alpha?$$

You should check this.

8. We need to make more definitions to be able to simplify what we will say later. These words are not new for most of you, as you already heard them in MAT223. We define the kernel of f as

$$\ker(f) = \{v \in V \text{ such that } f(v) = 0\}.$$

Here are some easy but important consequences of this definition.

(a) Show that $\ker f$ is a subspace of V.

(b) Show that if f is one-to-one, then $\ker(f) = \{0\}$.

(c) Show that if $\ker(f) = \{0\}$, then f is one-to-one.

(d) Show that if v is in $\ker(f)$, then $[v]_\alpha$ is in the nullspace of $[f]_\alpha^\beta$.

(e) Show that if $[v]_\alpha$ is in the nullspace of $[f]_\alpha^\beta$, then v is in $\ker(f)$.

9. We define the image (range) of f to be

$$\operatorname{im}(f) = \{w \in W \text{ such that there is a } v \in V \text{ with } f(v) = w\}.$$

(a) Show that $\operatorname{im}(f)$ is a subspace of W.

(b) Conclude that $\dim(\operatorname{im}(f)) \leq \dim W$.

(c) Show that $\operatorname{im}(f) = W$ iff f is onto (Notice that there is nothing to show here, this is actually definition of being onto).

(d) Show that $w \in \operatorname{im}(f)$ iff $[w]_\beta$ is in the column space of $[f]_\alpha^b eta$.

(e) Note that

$$\begin{aligned} w \in \operatorname{im}(f) &\iff f(v) = w \text{ for some } v \in V \\ &\iff f(c_1v_1 + \ldots + c_nv_n) = w \text{ for some } c_1, \ldots, c_n \in \mathbb{R} \\ &\iff c_1f(v_1) + \ldots + c_nf(v_n) = w \text{ for some } c_1, \ldots, c_n \in \mathbb{R} \end{aligned}$$

Conclude from this that

$$\operatorname{im}(f) = \operatorname{span}\{f(v_1), \ldots, f(v_n)\}.$$

10. Combining Q8d and Q8e, we can argue that

$$\dim(\ker(f)) = \operatorname{nullity}([f]_\alpha^\beta)$$

and from 9c, we can conclude that

$$\dim(\operatorname{im}(f)) = \operatorname{rank}([f]_\alpha^\beta).$$

In MAT223, you learned the Rank-Nullity Theorem which we state as the Dimension Theorem as follows

$$\dim(\ker(f)) + \dim(\operatorname{im}(f)) = \dim V.$$

11. Now, we take a step back and prove the following facts.

(a) Let $x_1, \ldots, x_k \in V$ and assume $\{f(x_1), \ldots, f(x_k)\}$ is a linearly independent subset of W. Show that $\{x_1, \ldots, x_k\}$ is a linearly independent subset of V.

(b) Let $\{x_1, \ldots, x_k\}$ be a linearly independent subset of V and assume f is one-to-one. Show that $\{f(x_1), \ldots, f(x_k)\}$ is a linearly independent subset of W.

(c) Show with a counterexample that if f is not one-to-one, then the consequence of part b is no longer true. **Hint.** For many such questions, the zero transformation (which takes all the vectors in V to zero in W) is a counterexample. Try it first.

(d) Conclude that $\{f(v_1), \ldots, f(v_n)\}$ is a linearly independent subset of W if and only if f is one-to-one.

12. Combine 9b, 9e and 11d to conclude that if f is one-to-one, then $\dim V \leq \dim W$. You can also use the Dimension Theorem to prove this. Do it both ways.

13. Show that the converse of the previous question is not true. That is, if $\dim V \leq \dim W$, then f does not have to be one-to-one. **Hint.** Zero transformation.

14. Use 9b, 9c and 9e; or the Dimension Theorem to conclude that if f is onto, then $\dim V \geq \dim W$.

15. Show that the converse of the previous question is not true. That is, if $\dim V \geq \dim W$, then f does not have to be onto. **Hint.** Zero transformation. once again!

16. We will now go back to matrix representations. Here is another very important formula whose proof is tedious but quite simple. If U is a third vector space with basis γ and if $g : W \to U$ is another linear transformation, prove the followings.

    (a) $g \circ f : V \to U$ is a linear transformation. (This is straightforward).

    (b)

    $$[g \circ f]_\alpha^\gamma = [g]_\beta^\gamma \cdot [f]_\alpha^\beta$$

    This just follows from some calculations. Construct your own example to see this in action. This is quite useful as we will see soon.

    Note. We are aware that there are some *tutoring companies* out there which sells notes to our students. You should know that these notes are freely available and these so called tutoring companies are scams. If you need help, we have more than 10 hours of office hours per week available for you with about 16 hours during midterm weeks. If you went to all these office hours and you still needed help, email us. These companies do not know anything about our midterms. A good preparation for our midterm is to understand the material here and the computational problems on the problem sets. Worry about your knowledge on the material, don't try to guess which questions can be on the midterm. You can not learn the material if somebody hands you the solutions without you trying them on your own. We care about you. Don't be scammed.

17. We say that a linear transformation is an isomorphism if it is one-to-one and onto (hence invertible). We say that V is isomorphic to W if there is an isomorphism from V to W.

    (a) Show that the linear transformation $T_\alpha : V \to \mathbb{R}^n$ defined by the formula

    $$T_\alpha(v) = [v]_\alpha$$

    is an isomorphism. This is not difficult, but it is very important. Make sure you understand how this works.

    (b) Conclude that if $\dim V = n$, then V is isomorphic to $\mathbb{R}^n$.

    (c) Show that if f is an isomorphism, then $f^{-1} : W \to V$ is also an isomorphism. **Hint.** To show that $f^-1(x + y) = f^{-1}(x) + f^{-1}(y)$, you can let $a = f^{-1}(x + y), b = f^{-1}x, f^{-1}(c) = y$. Now your purpose is to show that $a = b + c$. Compute $f(a)$ and $f(b + c)$. Then, use the fact that f is one-to-one.

(d) Conclude that if V is isomorphic to W, then W is isomorphic to V.

(e) Show that if f is an isomorphism and $g : W \to U$ is an isomorphism, then $g \circ f$ is also an isomorphism.

18. Conclude using 17b, 17d and 17e that if $\dim V = \dim W$, then V is isomorphic to W.

19. Use 12 and 14 to conclude that if V is isomorphic to W, then $\dim V = \dim W$.

20. Combine the previous two questions to conclude that V is isomorphic to W if and only if $\dim V = \dim W$. Notice that we did not construct an explicit isomorphism to prove this result, we used abstract linear algebra. If we want to construct an isomorphism, here is a recipe: send v_1 to w_1, v_2 to w_2 and so on until v_n to w_n. Show that this recipe actually defines an isomorphism. We can choose any other bases, and this recipe gives an isomorphism.

21. Be careful. We can have $\dim V = \dim W$ and a linear map between them which is not an isomorphism. **Hint.** Zero transformation once again!

22. Show that if $I : V \to V$ is the identity transformation on V then I is a linear transformation. Show that $[I]_\alpha^\alpha$ is the $n \times n$ identity matrix I_n.

23. Assume f is an isomorphism. Using the fact that $f^{-1} \circ f = I$, and using 16 with $f^{-1} : W \to V$ and $f : V \to W$ and with $\gamma = \alpha$, conclude

$$[f^{-1}]_\beta^\alpha \cdot [f]_\alpha^\beta = [I]_\alpha^\alpha$$

Using 22, conclude that

$$[f^{-1}]_\beta^\alpha = \left([f]_\alpha^\beta\right)^{-1}.$$

24. In 22, you showed that $[I]_\alpha^\alpha$ is the identity matrix. Show with a counterexample that if α' is another basis, then $[I]_\alpha^{\alpha'}$ is not the identity matrix. This matrix has a special name. It is called the change of basis matrix from α to α'. It is the matrix of the identity transformation with respect to bases α and α'.

25. Using 16, 22 and the fact that $I \circ I = I$, conclude

$$[I]_{\alpha'}^\alpha = \left([I]_\alpha^{\alpha'}\right)^{-1}$$

26. Now, we will need some subscripts. Let's denote by I_V the identity transformation on V and by I_W the identity transformation on W. Let α' and β' be bases for V and W respectively.

(a) Convince yourself that $f = I_W \circ f \circ I_V$.

(b) Conclude using 16 that

$$[f]_{\alpha'}^{\beta'} = [I]_\beta^{\beta'} \cdot [f]_\alpha^\beta \cdot [I]_{\alpha'}^\alpha$$

You just proved the change of basis formula.

27. **NOTE.** From now on, we will assume that $W = V$. So, we have a linear operator $f : V \to V$. (Operator is just a fancy name to say that the domain and the codomain are the same, a nice property of an operator is that we can compose it with itself, so it allows us to model some dynamical systems.)

28. Let β be a basis for V. Use 26 to conclude that

$$[f]_\alpha^\alpha = [I]_\beta^\alpha \cdot [f]_\beta^\beta \cdot [I]_\alpha^\beta$$

Use 25 to argue that

$$[f]_\alpha^\alpha = \left([I]_\alpha^\beta\right)^{-1} \cdot [f]_\beta^\beta \cdot [I]_\alpha^\beta$$

29. Suppose that there is a basis $\alpha = \{v_1, \ldots, v_n\}$ of V such that

$$f(v_1) = \lambda_1 v_1, \quad f(v_2) = \lambda_2 v_2, \quad \ldots \quad , f(v_n) = \lambda_n v_n$$

for some $\lambda_1, \ldots, \lambda_n$. Compute $[f]_\alpha^\alpha$.

30. If you actually computed $[f]_\alpha^\alpha$ in the previous question, you see that you have a diagonal matrix. If you didn't, do it now. This is important. The nice thing about diagonal matrices is that they are so much easier to make computations with. Here is an example. Compute the 7th power of the following two matrices. Measure time while doing it.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Can you compute the 100th powers? When a process is modeled by a linear operator and you want to understand what happens in the 100th step or 100th day, then you want to be able to compute these powers. When you have larger matrices, even wolframalpha can't save you. You need matrices which are as close as to a diagonal matrix. If there is a basis as in 29, you are lucky. You end up with a diagonal matrix. If there is no such basis, then we will see what happens at the end of the course.

31. Since we don't want to keep refering to 29, we should make a definition and give names to such vectors. Let $f : V \to V$ be a linear operator. A nonzero vector $v \in V$ is called an eigenvector of f if there is a λ such that $f(v) = \lambda v$. Such a λ is called an eigenvalue. Show that

$$E_\lambda = \{v \in V \quad : \quad f(v) = \lambda v\}$$

is a subspace of V.

32. Show that for any $v \in E_\lambda$, we have $f(v) \in E_\lambda$. Conclude from this that we can restrict our function f to E_λ and its image lies in E_λ:

$$f|_{E_\lambda} : E_\lambda \to E_\lambda$$

Pick a basis β for E_λ and show that $[f|_{E_\lambda}]_\beta^\beta$ is a diagonal matrix with λs on the diagonal.

33. Note that $f(v) = \lambda v$ if and only if $f(v) - \lambda v = 0$. This is equivalent to saying that $f(v) - (\lambda I)(v) = 0$, which is equivalent to saying that $(f - \lambda I)(v) = 0$. From this conclude that

$$E_\lambda = \ker(f - \lambda I).$$

34. Note that

$$\begin{aligned} f(v) = \lambda v &\iff [f(v)]_\alpha = [\lambda v]_\alpha \\ &\iff [f(v)]_\alpha = \lambda [v]_\alpha \\ &\iff [f]_\alpha^\alpha [v]_\alpha = \lambda [v]_\alpha \end{aligned}$$

From here, conclude that v is an eigenvalue of f if and only if $[v]_\alpha$ is an eigenvalue of $[f]_\alpha^\alpha$. You learned in MAT223 how to find eigenvalues and eigenvectors of matrices. So, finding eigenvalues and eigenvectors of operators is something you already learned. All you need to do is to pick a basis and write the corresponding matrix. Then, you are back in MAT223 and you can find the eigenvalues.

35. If you don't remember how to find eigenvalues of a matrix A, here is a reminder.

$$\begin{aligned} Ax = \lambda x &\iff Ax = \lambda I x \\ &\iff Ax - \lambda I x = 0 \\ &\iff (A - \lambda I)x = 0 \\ &\iff x \in \operatorname{null}(A - \lambda I) \end{aligned}$$

In MAT223, you learned that there is a nonzero vector in the nullspace of a matrix if and only if that matrix is not invertible. And one of the big theorems said that this is if and only if the determinant of this matrix is zero. Hence, we conclude that λ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$. Note that $\det(A - \lambda I)$ is a polynomial in λ, if you don't believe this, see it on a couple of examples to convince yourself.

36. There is a very important point here. Go back to 5 and convince yourself that that if we choose two different bases α and β, then $[f]_\alpha^\alpha$ and $[f]_\beta^\beta$ are different matrices. 28 tells you the relation between them. Here is a nice observation: Even though they are different matrices, they have the same characteristic polynomial! Here is how you can prove it. To make things simpler, let's make a definition here. Let A and B be two $n \times n$ matrices. We say that A is similar to B if there is an invertible matrix S such that $A = S^{-1}BS$.

    (a) Show that if A is similar to B, then B is similar to A. (You can show $B = SBS^{-1}$. Remember $S = S^{-1^{-1}}$).

    (b) Show that if A is similar to B, then for any $\lambda \in \mathbb{R}$, $A - \lambda I$ is similar to $B - \lambda I$. **Hint.** $X(Y + Z)W = XYW + XZW$.

    (c) Show that if A is similar to B, then $\det(A) = \det(B)$. **Hint.** $\det(XY) = \det(X)\det(Y)$.

    (d) Use the previous two parts to conclude that if A is similar to B, then they have the same characteristic polynomial.

    Use this exercise and 28 to conclude what we wanted to prove at the beginning of this exercise. Since it doesn't matter which basis we choose we define the characteristic polynomial of f to be the characteristic polynomial of $[f]_\alpha^\alpha$ for some basis α.

37. We say that a linear operator $f : V \to V$ is diagonalizable if there is a basis β of V consisting of eigenvectors of f. See 29 to understand why this name makes sense.

38. Suppose that $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator and suppose that T is diagonalizable. Let std be the standard basis of $\mathbb{R}^n$ and let α be a basis of $\mathbb{R}^n$ consisting of eigenvectors of T. Then, write the change of basis matrix $[I]_{\text{std}}^\alpha$ from std to α and the change of basis matrix $[I]_\alpha^{\text{std}}$ from α to std. Observe from 25 that these matrices are inverses of each other. What does $[I]_\alpha^{\text{std}}$ look like? What are its columns? Recall $[T]_\alpha^\alpha$ is a diagonal matrix and

$$[T]_\alpha^\alpha = [I]_\alpha^{\text{std}} \cdot [T]_{\text{std}}^{\text{std}} \cdot \left([I]_\alpha^{\text{std}}\right)^{-1}$$

This is how you diagonalize your matrix.

39. Happy studying! Come to office hours for questions and/or a chat. Good luck in your exam.