

## ASSIGNMENT 2: CHAPTERS 1 & 2

This document contains suggested solutions to a selection of *representative* problems (which are graded) from the current homework assignment—one, or two, from each subsection of chapter one in the textbook. They do not include every last detail, but should give you enough guidance to complete each problem fully on your own.

Even though the rest of the homework assignment are checked for completion, the grader will make an attempt at pointing out glaring faults in arguments whenever they arise. As such, some general comments on the completion of this assignment are provided at the end.

Below are a few conventions are used throughout grading your assignments, along with some ground rules for grading:

- (1) Up to 40 raw marks are given for completion of homework. The actual score given depends on the amount of assignment completed, and quality of the work attempted. This will be given in 5-mark increments. As a rough description for this raw score, the table below describes how this score will be assigned:

Mark(s)	Description
0	Did not complete anything in the assignment / did not hand it in.
10	A lighthearted attempt at a few questions are made.
20	Up to half of the questions have been attempted, with some effort put in.
30	More than three-fourths of the questions have been attempted, all with considerable progress.
40	Attempted every question assigned, and has made considerable progress in every question.

- (2) Marks available and breakdown are indicated at the beginning of every question, inside square brackets.
- (3) Naming convention: in the following section, question indexed as  $2.x.y$  should be read as “question  $y$  from chapter  $2.x$  in Lay’s textbook”.
- (4) If you wish to dispute grading on question(s), please hand in your assignment to the instructor on the next homework submission day—with your comments/disputes written around the question(s).

SUGGESTED SOLUTIONS

- 1.9.33** [6; 1 for defining the bases of linear transformation, 1 for stating/acknowledging the definition of standard matrix, 2 for realising that  $Be_j = j$ -th column of  $B$ , 1 for correct calculations throughout and subsequent conclusion, 1 for having all of the above]

*Proof.* It is not difficult to notice that the matrix of linear transformation is dependent on the choice of basis, denoted  $\mathcal{B}$ . In this case, let us deal with the canonical (standard) bases. Take a linear transformation  $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ . Let  $(e_1, \dots, e_n)$  denote the canonical basis in  $\mathbb{R}^n$ , and  $(e'_1, \dots, e'_m)$  denote the canonical basis in  $\mathbb{R}^m$ . Let  $A$  denote the standard matrix of linear transformation defined as

$$A := \begin{bmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{bmatrix} \in \mathcal{M}_{m \times n}$$

where  $\mathcal{M}_{m \times n}$  denotes the set of  $m \times n$  matrices. Suppose there exists another  $m \times n$  matrix  $B$  such that  $T(x) = Bx$ . It follows trivially that  $j$ -th column of  $A = T(e_j) = Be_j = j$ -th column of  $B$ . Hence, the matrix  $A$  and  $B$  share the same columns. This concludes the proof. ■

- 2.1.16** [10; 2 for each part—1 for T/F, 1 for correct justification]

- (a) **False.** Check the dimensions of the resulting matrix/vector!
- (b) **True.** Follows from definition of matrix multiplication.
- (c) **False.** Matrix multiplication is not commutative. Take  $A \in \mathcal{M}_{5 \times 6}$ ,  $B \in \mathcal{M}_{6 \times 4}$ ,  $C \in \mathcal{M}_{4 \times 6}$ . Then  $(AB)C \in \mathcal{M}_{5 \times 7}$ , but  $(AC)B$  is undefined. Even just with two matrices, such as  $B$  and  $C$  as defined here,  $BC \neq CB$ .
- (d) **False.**  $(AB)^\top = B^\top A^\top$ .
- (e) **True.** This is straightforward verification by playing around with the indices; note that for any matrix  $A$ ,  $A_{ij}^\top = A_{ji}$ ; by definition of matrix sum, for any two matrices  $A$  and  $B$  of same dimensions,  $(A + B)_{ij}^\top = (A + B)_{ji}$ . Fill in the rest of the details should you be interested.

- 2.2.21** [4; 1 for justifying the system  $A\vec{x} = \vec{b}$  always has a solution if  $A$  is invertible, 1 for correctly identifying the unique solution, 1 for explaining why the set of vectors is linearly independent, 1 for having all of the above]

*Proof.* For the first question, recall that if  $A$  is invertible, then, for all  $\vec{b}$  in the range,  $A\vec{x} = \vec{b}$  always admits a unique solution (which comes from using the fact that  $A^{-1}$  exists, so we can left multiply by  $A^{-1}$  to get  $A^{-1}A\vec{x} = A^{-1}\vec{b} \implies \vec{x} = A^{-1}\vec{b}$ ). Hence, in the case of  $A\vec{x} = \vec{0}$ , we have the unique solution  $\vec{x} = A^{-1}\vec{0} = \vec{0}$ . From an earlier proposition, we know that a set of vectors  $\{v_k\}_{k=1}^n$  is linearly independent iff the matrix equation  $A\vec{x} = \vec{0}$  (where the columns of  $A$  are  $\{v_k\}_{k=1}^n$ ) admits only the trivial solution. This concludes the proof. ■

- 2.2.22** [4; 1 for an attempt at using the hint, 1 for justifying one of the equivalent conditions in Theorem 4, 1 for correct explanation, 1 for having all of the above]

*Proof.* For the second question, note that, by Theorem 5, for every  $\vec{b} \in \mathbb{R}^n$ ,  $A\vec{x} = \vec{b}$  always admits a unique solution. Per the hint, by Theorem 4, this is equivalent to columns of  $A$  spanning  $\mathbb{R}^n$ .

Alternatively, any invertible matrix  $A \in \mathcal{M}_{n \times n}$  is row equivalent to  $I_n$  (identity matrix of size  $n \times n$ ; if you struggle to understand this, think about the rank of any invertible

matrix). Since  $I_n$  has a pivot in every row, per the hint, by Theorem 4, this is equivalent to columns of  $A$  spanning  $\mathbb{R}^n$ . This concludes the proof. ■

**2.3.33 + 2.3.34** [8; 4 for each question. For each part, 1 for justifying the map is invertible, 1 for finding the inverse of the given matrix, 1 for stating (1) or (2) for respective questions, 1 for having all of the above. If one fails to justify that the map is invertible, even if the answer is correct, maximum of 1 is given for each question]

For the first question, denote the linear transformation as  $T_1$ . Note that  $T_1 : \mathbb{R}^2 \mapsto \mathbb{R}^2$  can be represented by the vector function

$$T_1 \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \underbrace{\begin{bmatrix} -5 & 9 \\ 4 & -7 \end{bmatrix}}_{:=A_1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

It is a fact that a linear map is invertible iff it is bijective; it remains to check that a map is bijective. It is not difficult to see that

$$\begin{bmatrix} -5 & 9 \\ 4 & -7 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

which implies it is one-to-one and onto by Theorem 12 in §1.9. Hence,  $T_1$  is bijective, therefore invertible (alternatively, an equivalent method is to check if the matrix representation of linear transformation  $A_1$  is invertible, and justify this by a theorem in this section). Lastly, we know such inverse is given by  $T_1^{-1}(x) = A_1^{-1}x$ , where  $x \in \mathbb{R}^2$ . A simple calculation yields

$$A_1^{-1} = \frac{1}{35 - 36} \begin{bmatrix} -7 & -9 \\ -4 & -5 \end{bmatrix} = \begin{bmatrix} 7 & 9 \\ 4 & 5 \end{bmatrix}$$

Hence,

$$(1) \quad T_1^{-1} \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} 7 & 9 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 7x_1 + 9x_2 \\ 4x_1 + 5x_2 \end{bmatrix}$$

For the second question, denote the linear transformation as  $T_2$ . Note that  $T_2 : \mathbb{R}^2 \mapsto \mathbb{R}^2$  can be represented by the vector function

$$T_2 \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \underbrace{\begin{bmatrix} 6 & -8 \\ -5 & 7 \end{bmatrix}}_{:=A_2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Note that because  $\det(A_2) = 42 - (40) = 2 \neq 0$ ,  $A_2$  is invertible. By an earlier theorem,  $T_2$  is an invertible linear map. A quick calculation gives

$$A_2^{-1} = \frac{1}{42 - 40} \begin{bmatrix} 7 & 8 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} \frac{7}{2} & 4 \\ \frac{5}{2} & 3 \end{bmatrix}$$

Hence,

$$(2) \quad T_2^{-1} \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} \frac{7}{2} & 4 \\ \frac{5}{2} & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{7}{2}x_1 + 4x_2 \\ \frac{5}{2}x_1 + 3x_2 \end{bmatrix}$$

**2.8.22** [10; 2 for each part—1 for T/F, 1 for correct justification]

(a) **False.** This is only one of the conditions required for  $H \subset \mathbb{R}^n$  to be a subspace.

- (b) **True.** The set of all linear combinations of a collection of vectors is the span of the collection of vectors; it is not hard to prove/convince yourself that this is indeed a vector space. In particular, it is a vector subspace of  $\mathbb{R}^n$ .
- (c) **True.** If  $A \in \mathcal{M}_{m \times n}$ , then the function  $\Gamma(\vec{x}) = A\vec{x}$  represents the linear map  $\Gamma : \mathbb{R}^n \mapsto \mathbb{R}^m$ . By the nature of matrix/vector multiplication,  $x$  is of size  $n \times 1$ . Hence, the nullspace is defined as

$$\mathbf{N}(A) = \left\{ \vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}_{\mathbb{R}^m} \right\}$$

where  $\vec{0}_{\mathbb{R}^m}$  denotes the zero vector in  $\mathbb{R}^m$ . It is not difficult to verify that this is indeed a subspace of  $\mathbb{R}^n$ .

- (d) **False.** The column space of  $A$  can be the entire ambient space (of the range). As a counterexample, consider

$$\begin{bmatrix} 3 & 0 \\ 0 & 4 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

obviously, the columns of  $A$  span  $\mathbb{R}^2$ . As such,  $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbf{C}(A)$ , but this is obviously not a solution to the matrix equation above.

- (e) **False.** One must look at the corresponding (pivot) columns of  $A$ !

**2.9.18** [10; 2 for each part—1 for T/F, 1 for correct justification]

- (a) **True.** This is not difficult to prove at all. For those interested, here is the proof:

*Proof.* Take any (finite) vector space  $V$  with basis  $\mathcal{B} := \{v_1, v_2, \dots, v_n\}$ , and any  $x \in V$ . By definition of  $\mathcal{B}$ , we can write  $x = \sum_{k=1}^n a_k v_k$  for some set of scalars  $a_k$ . By way of contradiction, suppose this representation is not unique; then,  $x = \sum_{k=1}^n b_k v_k$ , for some different set of scalars  $b_k$ . Then,

$$\sum_{k=1}^n a_k v_k - \sum_{k=1}^n b_k v_k = \sum_{k=1}^n (a_k - b_k) v_k = x - x = 0$$

where  $0$  is the zero vector in  $V$ . By definition of linear independence,  $a_k - b_k = 0$  for all  $k = 1, 2, \dots, n$ , implying  $a_k = b_k$  for all  $k$ . Contradiction. ■

- (b) **True.** Such map is bijective, even if the vectors in  $H$  themselves may have more than  $p$  entries.
- (c) **False.** For any matrix  $A$ ,  $\dim(\mathbf{N}(A)) = \text{number of free variables in the reduced form of } A$ .
- (d) **True.** This is the consequence of rank-nullity theorem.
- (e) **True.** This is the consequence of the basis theorem.

**3.1.39** [4; 2 for each part—1 for T/F, 1 for correct justification]

- (a) **True.** The statement is not great, but it is essentially how (repeated) cofactor expansion works.
- (b) **False.** The cofactor is given by  $C_{ij} = (-1)^{i+j} \det(A_{ij})$ , where  $A_{ij}$  is the matrix  $A$  with  $i$ -th row and  $j$ -th column removed.

**3.1.40** [4; 2 for each part—1 for T/F, 1 for correct justification]

- (a) **False.** We can expand down any row or column and get same *determinant*. In other words, the cofactor expansions of  $\det(A)$  down a row and down a column are the same.
- (b) **False.** The determinant of a triangular matrix is the product of the diagonal entries.

**Extra Credit 1** The full proposition is as follows:

**Proposition 1.** Let  $A$  be any  $m \times n$  matrix. Then the following hold:

- (a) The *left multiplication*  $E_{ij}A$  corresponds to swapping rows  $i$  and  $j$  actually swaps rows.
- (b) The *left multiplication*  $E_{k(i)}A$  corresponds to multiplying row  $i$  by  $k \in \mathbb{R} \setminus \{0\}$ .
- (c) The *left multiplication*  $E_{i+k(j)}A$  corresponds to multiplying row  $j$  by  $k \in \mathbb{R}$ , and adding it onto row  $i$ .

where the elementary matrices are defined as follows:

$$E_{ij} := \begin{bmatrix} E_1 \\ \vdots \\ E_{i-1} \\ E_j \\ E_{i+1} \\ \vdots \\ E_{j-1} \\ E_i \\ E_{j+1} \\ \vdots \\ E_m \end{bmatrix}, \quad E_{k(i)} := \begin{bmatrix} E_1 \\ \vdots \\ E_{i-1} \\ kE_i \\ E_{i+1} \\ \vdots \\ E_m \end{bmatrix}, \quad E_{i+k(j)} := \begin{bmatrix} E_1 \\ \vdots \\ E_{i-1} \\ E_i + kE_j \\ E_{i+1} \\ \vdots \\ E_{j-1} \\ E_j \\ E_{j+1} \\ \vdots \\ E_m \end{bmatrix}$$

where all elementary matrices are of size  $m \times m$ .

*Proof.* Each part can be proven by matrix multiplication. For simplicity, I omit writing out the details of the multiplications: they are all essentially summations of products.

(a) Note that

$$E_{ij}A = \begin{bmatrix} E_1A \\ \vdots \\ E_{i-1}A \\ E_jA \\ E_{i+1}A \\ \vdots \\ E_{j-1}A \\ E_iA \\ E_{j+1}A \\ \vdots \\ E_mA \end{bmatrix}$$

and since  $E_i = (0, 0, \dots, 0, \underbrace{1}_{i\text{-th position}}, 0, \dots, 0)$ ,  $E_iA$  returns the  $i$ -th row of  $A$ . *Mutatis mutandis*,  $E_jA$  returns  $j$ -th row of  $A$ . This corresponds to row swaps between rows  $i$  and  $j$ .

(b) Again, for any nonzero  $k \in \mathbb{R}$ , we have

$$E_{k(i)}A = \begin{bmatrix} E_1A \\ \vdots \\ E_{i-1}A \\ kE_iA \\ E_{i+1}A \\ \vdots \\ E_mA \end{bmatrix}$$

every  $j$ -th entry is now the  $j$ -th row of  $A$ , with the exception of  $i$ : it is now premultiplied by the scalar  $k$ . For all  $k \in \mathbb{R} \setminus \{0\}$ , this operation is invertible:  $E_{\frac{1}{k}(i)}E_{k(i)}A = A$ .

(c) By the distributive property of matrix multiplication, we have

$$E_{i+k(j)}A = \begin{bmatrix} E_1A \\ \vdots \\ E_{i-1}A \\ (E_i + kE_j)A \\ E_{i+1}A \\ \vdots \\ E_{j-1}A \\ E_jA \\ E_{j+1}A \\ \vdots \\ E_mA \end{bmatrix} = \begin{bmatrix} E_1A \\ \vdots \\ E_{i-1}A \\ E_iA + kE_jA \\ E_{i+1}A \\ \vdots \\ E_{j-1}A \\ E_jA \\ E_{j+1}A \\ \vdots \\ E_mA \end{bmatrix}$$

from previous parts, we know  $E_iA$  corresponds to row  $i$  of matrix  $A$ , and  $kE_jA$  to  $k$  multiples of row  $j$  of matrix  $A$ .

Combining the three parts above concludes the proof. ■

**Extra Credit 2** We prove the following theorem:

**Theorem 2** (The Basis Theorem). Let  $H \subset \mathbb{R}^n$  be a nontrivial subspace of  $\mathbb{R}^n$ , with  $\dim(H) = k$ . Then, a collection of vectors  $V := \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$  is spanning iff it is linearly independent. Stated another way, the following are equivalent:

- (a)  $V$  is a basis of  $H$ .
- (b)  $V$  is a maximal linearly independent set.
- (c)  $V$  is a minimal spanning set of  $H$ .

*Proof.* By definition of a basis, (a) implies (b) and (c). To show (b) and (c) are equivalent, we will show (b) implies (a) and (c) implies (a).

Assume (b). By basis reduction theorem, we can remove some  $v_i$  to obtain a basis of  $H$  (try to prove this proposition). However, every basis of  $H$  must have length  $k$ , hence no vector needs to be removed from  $V$ . The set  $V$  is already a basis of  $H$ .

Assume (c). In any finite-dimensional vector space  $V$ , every linearly independent set of vectors can be extended to a basis of  $V$  (try prove this proposition).  $V$  can be extended as well. However, by the same argument, every basis of  $H$  must have length  $k$ , hence no vector needs to be added. As such,  $V$  is a basis of  $H$ .

As such, the statements above are equivalent. ■

*Alternative proof (from first principles).* First, we prove a useful lemma.

**Lemma 3.** Let  $H \subseteq \mathbb{R}^n$  be a subspace of dimension  $k$ . If  $W \subseteq H$  is a  $k$ -dimensional subspace of  $H$ , then  $W = H$ .

*Proof.* Fix a basis  $\{\vec{w}_1, \dots, \vec{w}_k\}$  for  $W$ . We know that  $\text{span}(\vec{w}_1, \dots, \vec{w}_k) \subseteq H$  because  $H$  is closed under addition and scaling. So, if we suppose  $\text{span}(\vec{w}_1, \dots, \vec{w}_k) = W \neq H$ , it must be the case that there is some vector  $\vec{w} \in H$  such that

$$a_1\vec{w}_1 + \dots + a_k\vec{w}_k = \vec{w}$$

has no solution for  $a_1, \dots, a_k \in \mathbb{R}$ . If this is the case, then one can never write

$$a_1\vec{w}_1 + \dots + a_k\vec{w}_k + a_{k+1}\vec{w} = \vec{0}$$

unless all of the  $a_i$  are 0. So,  $\{\vec{w}_1, \dots, \vec{w}_k, \vec{w}\}$  is linearly independent in  $H$ , which implies that

$$\dim(H) \geq k + 1 > k = \dim(H).$$

This is a contradiction. So, no such  $\vec{w} \in H$  can exist. That is,  $\text{span}(\vec{w}_1, \dots, \vec{w}_k) = H$ . Thus,  $W = H$ . ■

Back to the task at hand. Take  $\{\vec{v}_1, \dots, \vec{v}_k\} \subseteq H$  a linearly independent set. Then, if  $A = [\vec{v}_1 \dots \vec{v}_k]$ , the Rank-Nullity Theorem tells us

$$\begin{aligned} k &= \dim \mathbf{C}(A) + \dim \mathbf{N}(A) \\ &= \dim \mathbf{C}(A) + 0 \end{aligned}$$

whence  $k = \dim \mathbf{C}(A)$ . So,  $\mathbf{C}(A) = H$ . So,  $\text{span}(\vec{v}_1, \dots, \vec{v}_k) = H$ , by Lemma 3. This proves the first statement.

For the second statement, suppose that  $\{\vec{v}_1, \dots, \vec{v}_k\}$  is a spanning set for  $H$ , so that  $\text{span}(\vec{v}_1, \dots, \vec{v}_k) = H$ . Then, consider the matrix  $A$  as before. We know that  $\dim \mathbf{C}(A) = k$ . So, by the Rank-Nullity Theorem:

$$\begin{aligned} k &= \dim \mathbf{C}(A) + \dim \mathbf{N}(A) \\ &= k + \dim \mathbf{N}(A) \end{aligned}$$

from which it follows that  $\dim \mathbf{N}(A) = 0$ . So,  $\mathbf{N}(A) = \{\vec{0}\}$ . So, there are no nontrivial linear combinations of  $\vec{v}_1, \dots, \vec{v}_k$ , implying the set of vectors is linearly independent. ■

**Remark 4.** The point of this extra credit question is to demonstrate that when one verifies whether a set of vectors is a basis of some subspace or not, one either checks if it spans the subspace, or if it is linearly independent.

Also, the theorem holds when vector space  $V$  is a finite-dimensional, which guarantees that the dimension of the subspace is finite. In other words, if  $V$  is finite, and  $H \subset V$ , then  $\dim(H) \leq \dim(V)$ .

## GENERAL COMMENTS

[Developing...]

The point of studying mathematics is to be able to come up with *simple, nontrivial examples*, from which one can think about the fundamental principles underlying them. Some of these T/F questions help one in doing exactly that. Some of these proofs-type questions also test one's understanding of the fundamentals—including, but not limited to, definitions, basic propositions, some theorems and their consequences, and so on.

Here are some general comments for each question (some of which familiar from the previous homework):

**1.9.33**

**2.1.16**

**2.2.21**

**2.2.22**

**2.3.33 + 2.3.34**

**2.8.22**

**2.9.18**

**3.1.39**

**3.1.40**

**Extra Credit 1**

**Extra Credit 2**