

ASSIGNMENT 5: CHAPTERS 5 & 6

This document contains suggested solutions to a selection of *representative* problems (which are graded) from the current homework assignment—a couple from each subsection assigned. 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 25 marks will be provided for successful completion of the assignment. As always, the actual score given depends on the amount of assignment completed, and quality of the work attempted.

Mark(s)	Description
0	Did not complete anything in the assignment / did not hand it in.
5	A lighthearted attempt at a few questions are made.
10	Attempts at almost half of the questions were provided—most of which lighthearted.
15	Up to half of the questions have been attempted, with some effort put in.
20	More than three-fourths of the questions have been attempted, all with considerable progress.
25	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 5.x.y should be read as “question y from chapter 5.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

- As Alekos noted, the notation $\mathcal{P}_n(\mathbb{R})$ is used to denote the space of real polynomials of degree $\leq n$ (instead of \mathbb{P}_n).
- Instead of Lay's notation $M_{m \times n}$ for the (vector) space of $m \times n$ matrices with entries in \mathbb{R} , the notation $\mathcal{M}_{m \times n}(\mathbb{R})$ is used throughout here. For square matrices, we will use the simplified notation $\mathcal{M}_n(\mathbb{R})$ to denote the set of $n \times n$ matrices with entries over \mathbb{R} .

5.4.6 [9; see below for breakdown]

For this question, let the map be denoted

$$\begin{aligned} \mathcal{T} &: \mathcal{P}_2(\mathbb{R}) \longrightarrow \mathcal{P}_4(\mathbb{R}) \\ p(t) &\longmapsto p(t) + t^2 p(t) \end{aligned}$$

- (a) [3; 1 for any attempt, 1 for correct calculations and final answer, 1 for having all of the above]

Let $a(t) = 2 - t + t^2$. Then,

$$\begin{aligned} \mathcal{T}(a(t)) &= (2 - t + t^2) + t^2(2 - t + t^2) \\ &= 2 - t + 3t^2 - t^3 + t^4 \end{aligned}$$

- (b) [3; 1 for any attempt, 1 for correct verifications of linearity and homogeneity, 1 for having all of the above]

Proof. This should be standard at this point. For all $a, b \in \mathbb{R}$ and $p, q \in \mathcal{P}_2(\mathbb{R})$ (dropping the argument on the polynomials for notational simplicity),

$$\begin{aligned} \mathcal{T}(ap + bq) &= (ap + bq) + t^2(ap + bq) \\ &= ap + (bq + t^2ap) + t^2bq && \text{associativity} \\ &= ap + (t^2ap + bq) + t^2bq && \text{commutativity} \\ &= (ap + t^2ap) + (bq + t^2bq) && \text{commutativity} \\ &= a(p + t^2p) + b(q + t^2q) && \text{linearity} \\ &= a\mathcal{T}(p) + b\mathcal{T}(q) \end{aligned}$$

as desired. ■

- (c) [3; 1 for any attempt at calculating images of the vectors in basis, 1 for having all correct calculations, 1 for having all of the above]

Note that images of the standard basis of $\mathcal{P}_2(\mathbb{R})$ are

$$\mathcal{T}(1) = 1 + t^2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathcal{T}(t) = t + t^3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathcal{T}(t^2) = t^2 + t^4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

the matrix of linear transformation \mathcal{T} is then

$$M = \begin{bmatrix} 1 & & \\ & 1 & \\ 1 & & 1 \\ & 1 & \\ 2 & & 1 \end{bmatrix}$$

5.5.3 [8; 1 for any attempt, 1 for setting up the characteristic polynomial correctly, 1 for each correct eigenvalue—up to 2 available, 1 for attempting to find eigenvectors, 1 for each correct eigenvector—up to 2 available, 1 for having all of the above

Let $Q = \begin{bmatrix} 1 & 5 \\ -2 & 3 \end{bmatrix}$ act on vectors in \mathbb{C}^2 . Then, to diagonalise it, we find the characteristic polynomial

$$\det(Q - \lambda I) = (1 - \lambda)(3 - \lambda) + 10 = 0$$

which gives the equation $\lambda^2 - 4\lambda + 13 = 0$, implying the solutions are

$$\lambda_i = 2 \pm 3i, \quad i = 1, 2$$

where each i takes different sign on the imaginary part. We know the definition of eigenspace is equivalent to the nullspace of $(Q - \lambda I)$ (if this puzzles you, write down the definition of an eigenspace, and manipulate the equation $Qv = \lambda v$ in a way it makes sense). Hence, calculating each gives

$$\begin{bmatrix} 1 - 3i \\ 2 \end{bmatrix} \rightsquigarrow (2 + 3i)$$

$$\begin{bmatrix} 1 + 3i \\ 2 \end{bmatrix} \rightsquigarrow (2 - 3i)$$

where \rightsquigarrow denotes the relation of correspondence between eigenvector and eigenvalue. If you want to calculate the inverse of the matrix formed by eigenvectors above, you will find the diagonalised form

$$Q = \begin{bmatrix} 1 + 3i & 1 - 3i \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 2 - 3i & \\ & 2 + 3i \end{bmatrix} \begin{bmatrix} -\frac{i}{6} & \frac{1}{4} + \frac{i}{12} \\ \frac{i}{6} & \frac{1}{4} - \frac{i}{12} \end{bmatrix}$$

(the decomposition is completely optional to write out)

6.1.19 [10; 2 for each part—1 for T/F, 1 for correct justification]

- (a) **True.** By definition.
- (b) **True.** By commutativity of operations in \mathbb{R} .
- (c) **True.** Take the norm in \mathbb{R}^n ; we then have

$$\|u + v\|^2 = \langle u + v, u + v \rangle = \|u\|^2 + 2\langle u, v \rangle + \|v\|^2$$

$$\|u - v\|^2 = \langle u - v, u - v \rangle = \|u\|^2 - 2\langle u, v \rangle + \|v\|^2$$

and setting them equal gives $4\langle u, v \rangle = 0$.

- (d) **False.** A counterexample is as follows: let

$$A = \begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix}, \quad w = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and let $v = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \in C(A)$ (and, obviously, $w \in N(A)$). But observe $\langle v, w \rangle = 2 + 1 = 3 \neq 0$.

- (e) **True.** This is a part of a larger proposition, which is not difficult at all to prove.

Proposition 1. Let W be a subspace of \mathbb{R}^n spanned by p vectors, ie. $W = \text{span}\{v_1, v_2, \dots, v_p\}$. Then, $x \in W^\perp$ iff $x \perp w_i$ for all $i = 1, 2, \dots, p$.

But, for our purpose, we prove (\Leftarrow) direction. We have, for all $w \in W$ (which can be written as a linear combination of $\{v_i\}_{i=1}^p$),

$$\begin{aligned}\langle x, w \rangle &= \left\langle x, \sum_{i=1}^p c_i v_i \right\rangle \\ &= \sum_{i=1}^p \langle x, c_i v_i \rangle && \text{linearity} \\ &= \sum_{i=1}^p c_i \langle x, v_i \rangle && \text{homogeneity} \\ &= \sum_{i=1}^p c_i \times 0 = 0\end{aligned}$$

where the penultimate equality follows by assumption.

6.2.24 [10; 2 for each part—1 for T/F, 1 for correct justification]

- (a) **False.** Orthogonality implies linear independence, since $\langle u, v \rangle = 0$.
- (b) **False.** The vector(s) in the set does not need to be of unit length.
- (c) **True.** Isometries in \mathbb{R}^n preserve lengths, and elements of the orthogonal group are exactly those. The reason for this is as follows: let Q be an orthonormal matrix; then, for some vector v ,

$$v^T v = (Qv)^T (Qv) = v^T Q^T Q v = v^T v$$

by properties of transpose and orthonormal matrix.

- (d) **True.** By definition.
- (e) **True.** Columns are linearly independent, so such matrix is invertible by invertible matrix theorem.

6.3.11 [7; 1 for any attempt, 1 for showing correctly that $\{v_1, v_2\}$ is orthogonal, 1 for showing correctly that $\{v_1, v_2\}$ is linearly independent, 1 for stating correctly the orthogonal projection, 1 for attempting to calculate orthogonal projection, 1 for correct answer, 1 for having all of the above]

Observe that $\langle v_1, v_2 \rangle = 0$, so $\{v_1, v_2\}$ is an orthogonal set. We know that $W = \text{span}\{v_1, v_2\}$, and v_1, v_2 are linearly independent because both are nonzero. Hence, $\{v_1, v_2\}$ is an orthogonal basis for W .

We now know the answer is the orthogonal projection of x onto subspace W , denoted $\text{Proj}_W x$. The calculation is much less interesting:

$$\begin{aligned}\text{Proj}_W x &= \frac{\langle x, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \frac{\langle x, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 \\ &= \frac{1}{2} v_1 + \frac{3}{2} v_2 = \begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}\end{aligned}$$

6.4.17 [6; 2 for each part—1 for T/F, 1 for correct justification]

- (a) **False.** $c = 0$ is not what we want.
- (b) **True.** This is by definition.
- (c) **True.** Use $Q^T = Q^{-1}$.

6.7.16 [5; 1 for any attempt, 1 for using the fact that $\{u, v\}$ is orthonormal, 2 for correct calculations throughout, 1 for subsequently making the correct conclusion]

The verification is straightforward: for such orthonormal basis $\{u, v\}$, $\langle u, v \rangle = 0$. Then,
 $\|u - v\| = \sqrt{\langle u - v, u - v \rangle} = \sqrt{\|u\|^2 - 2\langle u, v \rangle + \|v\|^2} = \sqrt{1 - 2 \times 0 + 1} = \sqrt{2}$.

6.7.17 [5; 1 for any attempt, 1 for using linearity, 2 for correct calculations throughout, 1 for subsequently making the correct conclusion]

Starting from the RHS, we have $\|u + v\|^2 - \|u - v\|^2 = \langle u + v, u + v \rangle - \langle u - v, u - v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle - (\langle u, u \rangle + \langle u, -v \rangle + \langle -u, v \rangle + \langle v, v \rangle) \stackrel{\text{by linearity}}{=} 4\langle u, v \rangle$. As such, we have the relation $\langle u, v \rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$ by multiplying both sides by the multiplicative inverse of 4 (which exists because it is nonzero).

6.7.18 [5; 1 for any attempt, 1 for recognising linearity/homogeneity of inner products, 2 for correct calculations throughout, 1 for subsequently making the correct conclusion]

This is the **parallelogram identity in inner product spaces**, and is incredibly useful in a lot of contexts (just as useful as triangle and Cauchy-Schwarz inequalities). The proof is deceptively simple: from LHS, we have $\|u + v\|^2 + \|u - v\|^2 = \langle u + v, u + v \rangle + \langle u - v, u - v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle + (\langle u, u \rangle + \langle u, -v \rangle + \langle -u, v \rangle + \langle v, v \rangle) \stackrel{\text{by linearity}}{=} 2\|u\|^2 + 2\langle u, v \rangle - 2\langle u, v \rangle + 2\|v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

6.7.22 [5; 1 for any attempt, 1 for correct setup, 1 for some attempt at integration, 1 for correct answer, 1 for having all of the above]

Let $f(t) = 5t - 3$ and $g(t) = t^3 - t^2$. Then, by the inner product defined in the question, we have

$$\begin{aligned} \langle f, g \rangle &= \int_0^1 (5t - 3)(t^3 - t^2) dt \\ &= \int_0^1 (5t^4 - 8t^3 + 3t^2) dt \\ &= 1 - 2 + 1 = 0 \end{aligned}$$

6.7.24 [5; 1 for any attempt, 1 for correct setup, 1 for some attempt at integration, 1 for correct answer, 1 for having all of the above]

By the inner product defined in the question, we have

$$\begin{aligned} \|g\| &= \sqrt{\langle g, g \rangle} = \sqrt{\int_0^1 (t^3 - t^2)^2 dt} \\ &= \sqrt{\frac{1}{7} - \frac{1}{3} + \frac{1}{5}} = \frac{1}{\sqrt{105}} \end{aligned}$$

Below are the proofs of propositions laid out in the extra credit questions. These will be graded far more harshly than the questions from above. Marks are awarded for recognising overarching ideas in the propositions/theorems, and much less so for attempts.

Extra Credit 1 For some inner product space, choose an arbitrary set of basis $U = \{u_1, u_2, \dots, u_n\}$. Then, we construct the set of orthogonal vectors by projecting each vector onto the subspace spanned by the previous orthogonal vectors. In other words,

$$\begin{aligned} v_1 &= u_1 \\ v_2 &= u_2 - \text{Proj}_{v_1} u_2 \\ v_3 &= u_3 - \text{Proj}_{\text{span}\{v_1, v_2\}} u_3 \\ v_4 &= u_4 - \text{Proj}_{\text{span}\{v_1, v_2, v_3\}} u_4 \\ &\vdots \\ v_n &= u_n - \text{Proj}_{\text{span}\{v_1, v_2, v_3, \dots, v_{n-1}\}} u_n \end{aligned}$$

which can be simplified to

$$\begin{aligned} v_1 &= u_1 \\ v_2 &= u_2 - \frac{\langle v_1, u_2 \rangle}{\langle v_1, v_1 \rangle} v_1 \\ v_3 &= u_3 - \frac{\langle v_1, u_3 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle v_2, u_3 \rangle}{\langle v_2, v_2 \rangle} v_2 \\ v_4 &= u_4 - \frac{\langle v_1, u_4 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle v_2, u_4 \rangle}{\langle v_2, v_2 \rangle} v_2 - \frac{\langle v_3, u_4 \rangle}{\langle v_3, v_3 \rangle} v_3 \\ &\vdots \\ v_n &= u_n - \left(\sum_{i=1}^{n-1} \frac{\langle v_i, u_n \rangle}{\langle v_i, v_i \rangle} v_i \right) \end{aligned}$$

Extra Credit 2 Given the vector space $\mathcal{C}([0, 1], \mathbb{R})$ consisting of continuous functions $f : [0, 1] \rightarrow \mathbb{R}$, define an inner product by

$$(1) \quad \langle f, g \rangle = \int_0^1 f(t)g(t) dt$$

Before we answer the questions, let us rigorously verify that (1) is indeed an inner product in $\mathcal{C}([0, 1], \mathbb{R})$.

Proof. $\langle f, g \rangle = \langle g, f \rangle$, $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$ and $\langle cf, g \rangle = c \langle f, g \rangle$ are all trivially true $\forall f, g, h \in \mathcal{C}([0, 1], \mathbb{R})$ and $c \in \mathbb{R}$, using linearity of the Riemann integral, as well as commutativity and associativity of addition over \mathbb{R} . The much harder one is that $\langle f, f \rangle \geq 0$ under the same qualifiers from before. Of course, if f is the constant zero function, then the inner product gives zero. The more interesting case is when f is a nonzero continuous function. We now use the fact that f is continuous: for all $\varepsilon > 0$, there exists $\delta > 0$ s.t. $|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$. Renaming functions if you want (ie. let $g(x) = f(x)^2$), we can say that $\forall \varepsilon > 0, \exists \delta > 0$ s.t. $|x - y| < \delta \implies |f(x)^2 - f(y)^2| < \varepsilon$ (since $f(x)^2$ is still continuous as the composition of two continuous functions). In particular, this implies

$$(2) \quad |x - y| < \delta \implies f(x)^2 - f(y)^2 > -\varepsilon$$

We now want to approximate $\langle f, f \rangle$ by a quantity independent of the choice of x . Since the choice of ε is arbitrary, let $\varepsilon = \frac{f(y)^2}{2}$ for some $y \in (0, 1)$, and (2) reads $|x - y| < \delta \implies f(x)^2 > \frac{f(y)^2}{2}$. Then, we have

$$\begin{aligned} \langle f, f \rangle &= \int_0^1 f(x)^2 \, dx \\ &\geq \int_{y-\varepsilon}^{y+\varepsilon} f(x)^2 \, dx \\ &> \int_{y-\varepsilon}^{y+\varepsilon} \frac{f(y)^2}{2} \, dx \\ &= 2\varepsilon \frac{f(y)^2}{2} = \varepsilon f(y)^2 > 0 \end{aligned}$$

hence the property is verified. ■

In addition, the \mathcal{L}^2 -norm

$$(3) \quad \|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int_0^1 f(x)^2 \, dx}, \quad \forall f \in \mathcal{C}([0, 1], \mathbb{R})$$

is the norm in the space of \mathcal{L}^2 functions¹. The canonical way of defining this is by defining the Lebesgue integral (which agrees with the Riemann integral on the real line, hence (3) makes sense), and defining a measure λ (called the Lebesgue measure). Regardless of the technicalities, the quantity (3) still defines a norm on the space of continuous functions.

For the first part, we want the integrand to be periodic on $[0, 1]$. Trigonometric functions are periodic, and one can easily check that for $f(x) = \sin(\pi x)$ and $g(x) = \sin(2\pi x)$, $\int_0^1 f(x)g(x) \, dx = 0$. For another, if you want to find examples of polynomials that are orthogonal, #6.7.22 provides an example, for $f(x) = 5x - 3$ and $g(x) = x^3 - x^2$.

With a bit of knowledge in calculus, one can use the Taylor series expansions for $\sin x$ or $\cos x$ to see that trigonometric functions should be in our integrand. This is the basis of Fourier analysis, with which some intuitions on such integrals are available instantly. A possible example is $f(x) = \sqrt{2} \sin(\pi x)$ (a non-contrived one, that is). There are endless examples that are of the contrived nature—for instance, one can always rescale $f(x) = x$ to obtain $\tilde{f}(x) = \sqrt{3}\pi x$, and observe that $\|\tilde{f}\| = \pi$. A correct example (along with correct verifications) warrants the full marks.

Remarks #6.7.19 is an interesting problem. The problem can be solved with a few calculations:

$$\begin{aligned} |\langle u, v \rangle| &= \left| \begin{bmatrix} \sqrt{ab} \\ \sqrt{ab} \end{bmatrix} \right| = 2\sqrt{ab} \\ \|u\| &= \|v\| = \sqrt{a+b} \end{aligned}$$

and by Cauchy-Schwarz, we have $2\sqrt{ab} \leq a+b \implies \sqrt{ab} \leq \frac{a+b}{2}$. This is the famous *AM-GM inequality*, which is closely related to Jensen's inequality—in the sense that AM-GM can be proven from it. These inequalities are extremely useful in practical terms.

If you choose to study more mathematics, one of the most useful (and oddly mysterious) tools is Fourier analysis. The basic motivation for Fourier analysis is we want to represent

¹Remark: \mathcal{L}^2 is complete and is endowed with an inner product, making it a Hilbert space.

functions by an infinite sum (or series) of elementary functions. The way to go about it (in the classical sense, due to Fourier) is to choose an *orthonormal basis*, which is

$$(4) \quad \begin{aligned} \varphi_0(x) &= \frac{1}{\sqrt{2\pi}} \\ \varphi_{2n-1}(x) &= \frac{\cos nx}{\sqrt{\pi}} \\ \varphi_{2n}(x) &= \frac{\sin nx}{\sqrt{\pi}} \end{aligned}$$

and ask ourselves *in what sense does the sequence of functions converge to the original function*. This is a part of a general approach called *approximate identities*, and the validity of these approaches are not always easy to prove. It is, by and large, very difficult to prove the modes of convergence of these sequences of functions, and—if we can—the methods are nontrivial at best. Insight from linear algebra gets us a long way in doing Fourier analysis: *understanding* discrete Fourier transform, continuous linear functionals in Schwartz spaces, and analysis of smooth functions (and so on) all require a strong working knowledge of linear algebra.

One can never learn enough linear algebra in one's career as a mathematician, as it is one of the most fundamental tools we have to simplify the problems we have at hand.

GENERAL COMMENTS

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.

Some general comments for the homework (as a whole):

- Most people did an excellent job providing counterexamples, and I was pleasantly surprised by the growing number of people who attempted to justify their responses to T/F questions.
- To those who did the extra credit question(s): they were all valiant attempts. There were good ideas in those responses, and it was good seeing that you know what needs to be proven.

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

5.4.6

5.5.3

6.1.19

6.2.24

6.3.11

6.4.17

6.7.16

6.7.17

6.7.18

6.7.22

6.7.24

Extra Credit 1

Extra Credit 2