Exercise Guide for *Linear Algebra* by Peter Lax

Muthu Chidambaram

Last Updated: June 8, 2019

Contents

About

> *"A modern mathematical proof is not very different from a modern machine, or a modern test setup: the simple fundamental principles are hidden and almost invisible under a mass of technical details."*
> - Hermann Weyl

These notes contain short summaries of (my) proof ideas for exercises and some theorems from the book *Linear Algebra* by Peter Lax. I have tried to make the summaries as brief as possible; sometimes only one direction of proof, one line, or even one equation. My goal was to include enough information in the summaries so that someone reading would be able to reconstruct a full proof with all the details if necessary. However, these summaries are tuned to my own personal context, and as such I'm sure mileage will vary. I would greatly appreciate any feedback/fixes.

Also, I like when people include (what they presume to be) relevant quotes in their notes, so I have to ask you to forgive my haughtiness in starting these notes with a quote from Hermann Weyl.

1 Fundamentals

1.1 Exercise 1

$x + z = x = x + z' \implies z = z'$.

1.2 Exercise 2

$0x + x = (0 + 1)x = x$.

1.3 Exercise 3

Coefficients can be represented as row vectors.

1.4 Exercise 4

Function can be represented as row vector by letting $a_i = f(s_i)$ for each $s_i \in S$.

1.5 Exercise 5

Follows from exercises 3 and 4.

1.6 Exercise 6

$y_1 + z_1 + y_2 + z_2 = (y_1 + y_2) + (z_1 + z_2)$ and $k(y_1 + z_1) = ky_1 + kz_1$.

1.7 Exercise 7

$a \in Y \cap Z \implies ka \in Y, ka \in Z \implies ka \in Y \cap Z$.

1.8 Exercise 8

$k0 = 0, 0 + 0 = 0$.

1.9 Exercise 9

If S contains x_i then it must contain kx_i.

1.10 Exercise 10

If $x_i = 0$, k_i can be anything.

1.11 Exercise 11

$x = \sum_{i=1}^{m} \sum_{j=1}^{\dim Y_i} y_j^{(i)}$.

1.12 Exercise 12

Complete basis for W to U and V. Use W basis vectors and additional U and V basis vectors to get $\dim X = \dim U - \dim W + \dim V - \dim W + \dim W$.

1.13 Exercise 13

Send i^{th} basis vector to e_i, where e_i is vector of all zeroes except a one in the i^{th} place. Can permute mapping to get different isomorphisms.

1.14 Exercise 14

$x_1 - x_2 + x_2 - x_3 = x_1 - x_3$.

1.15 Exercise 15

$x' = x + z_x, y' = y + z_y \implies x' + y' = x + y + (z_x + z_y)$.

1.16 Exercise 16

$x \in X_1 \bigoplus X_2 \implies x = (x_1, x_2) = (x_1, 0) + (0, x_2)$.

1.17 Exercise 17

Construct a basis for X from Y: $y_1, ..., y_j, x_{j+1}, ..., x_n$.
Then $X/Y = \text{span}\{x_{j+1}, ..., x_n\}$.

2 Duality

Theorem 1

$x = \sum_{i=1}^{n} a_i x_i \implies k_i(x) = a_i$.

2.1 Exercise 1

$l_1, l_2 \in Y^\perp \implies l_1(y) + l_2(y) = 0 = (l_1 + l_2)(y)$.

2.2 Exercise 2

$\forall \xi \in Y^{\perp\perp} \implies \forall l \in Y^\perp,\ \xi(l) = 0 = l(y)\ \forall y \in Y$.

3 Linear Mappings

3.1 Exercise 1

(a) $x \in X \implies x = \sum_{i=1}^{n} k_i x_i \implies T(x) = \sum_{i=1}^{n} k_i T(x_i) \in U$.

(b) $T(x), T(y) \in U \implies T(x+y) \in U \implies x + y \in X$.

Theorem 1

$x \in X, y \in N_T \implies T(x+y) = T(x) + T(y) = T(x)$.

3.2 Exercise 2

(a) Differentiation constant and sum rules imply linearity, and multiplication by s is distributive. Take $p(s) = 1$ to see that $ST \neq TS$.

(b) Rotation by 90 degrees amounts to swapping and negating coordinates, which is linear. Take $p = (1, 1, 0)$ to see that $ST \neq TS$.

3.3 Exercise 3

(i) $T^{-1}(T(a+b)) = T^{-1}(T(a)+T(b)) = a+b = T^{-1}(T(a))+T^{-1}(T(b))$.

(ii) Composition of isomorphisms is an isomorphism, hence ST is invertible.

3.4 Exercise 4

(i) Let $T : X \to U$, $S : U \to V$ and $l_v \in V'$. Then $(ST)'(l_v) = l_v(ST) = (l_v S)T = (S' l_v)T = T'S' l_v$, since $S' l_v \in U'$.

(ii) Follows from linearity of transpose (definition).

(iii) Let $T : X \to U$ be an isomorphism. Then $l_x = l_u T \implies l_x T^{-1} = l_u$ for $l_u \in U'$, $l_x \in X'$.

3.5 Exercise 5

$T''(l_{x'}) = l_{x'} T'$ where $l_{x'} \in X''$ and $l_{x'} T' \in U''$. Since we can identify elements in X'' and U'' with elements in X and U respectively, we have that T'' assigns elements of U to X.

Theorem 2'

Since $T' : U' \to X'$ we have $l_u \in N_{T'} \implies T'(l_u) = l_u T = 0$. $N_{T'}^{\perp}$ consists of elements $l_{u'} | l_{u'}(l_u) = 0$. From $l_u T x = 0$ we have that each $l_{u'}$ is identified with a $u \in R_T$.

3.6 Exercise 6

The first two elements of x are already 0 after applying P, so $P^2 = P$. Linearity follows from linearity of vector addition.

3.7 Exercise 7

P is linear since function addition is linear. $P^2 f = \frac{f(x)+f(-x)}{4} + \frac{f(x)+f(-x)}{4} = Pf$.

4 Matrices

4.1 Exercise 1

$(P+T)_{ij} = ((P+T)e_j)_i = (Pe_j + Te_j)_i = P_{ij} + T_{ij}$.

4.2 Exercise 2

Represent A as a column of row vectors A_i and B as a row of column vectors B_i. Denote blocks by parenthesized subscripts. Then the first block of AB looks like:

$$
\begin{aligned}
(AB)_{(11)} &= \begin{pmatrix} A_1 B_1 & & \\ & \ddots & \\ & & A_k B_k \end{pmatrix} \\
&= \begin{pmatrix} A_{1,:(k+1)} B_{1,:(k+1)} & & \\ & \ddots & \\ & & A_{k,:(k+1)} B_{k,:(k+1)} \end{pmatrix} \\
&+ \begin{pmatrix} A_{1,(k+1):} B_{1,(k+1):} & & \\ & \ddots & \\ & & A_{k,(k+1):} B_{k,(k+1):} \end{pmatrix} \\
&= A_{(11)} B_{(11)} + A_{(12)} B_{(21)}
\end{aligned}
$$

Where:

$$
\begin{aligned}
A_{i,:(k+1)} B_{i,:(k+1)} &= \sum_{j=1}^{k} A_{i,j} B_{i,j} \\
A_{i,(k+1):} B_{i,(k+1):} &= \sum_{j=k+1}^{n} A_{i,j} B_{i,j}
\end{aligned}
$$

The rest follow similarly.

5 Determinant and Trace

5.1 Exercise 1

(a) The discriminant already has ordered versions of all the (i, j) difference terms. Applying a permutation only changes the signs of some of the difference terms, hence $\sigma(p) = 1, -1$.

(b) $\sigma(p_1 \circ p_2) = \text{sign}(P(p_1 \circ p_2(x_1, ..., x_n))) = \sigma(p_1)\text{sign}(P(p_2(x_1, ..., x_n)))$.

5.2 Exercise 2

(c) A transposition swaps two indices, and hence flips the sign of their associated difference term in the discriminant.

(d) If $p(i) = j$, then we can start with the permutation $(i\ j)$. Next, if $p(j) = k$, we can compose with $(i\ k)$ to get $(i\ k) \circ (i\ j)$. We can do this until we have completely reconstructed the permutation using transpositions.

5.3 Exercise 3

By starting with a different i in Exercise 2 (d), we can obtain a different decomposition of transpositions. However, the parity of the decomposition must be the same, as otherwise $\sigma(p)$ will take on two different values for the same p.

5.4 Exercise 4

(Property II): Each term in $D(a_1, ..., a_n)$ contains exactly one element from each of the a_i. Thus, scaling any of the a_i by k scales the entire determinant by k. Similar logic for vector addition.

(Property III): The only non-zero term in $D(e_1, ..., e_n)$ is associated with the identity permutation, hence $D(e_1, ..., e_n) = 1$.

(Property IV): Swapping two arguments is the same as applying a transposition to each of the terms in $D(a_1, ..., a_n)$, which flips the sign of D.

5.5 Exercise 5

Suppose $a_1 = a_2$. Then:

$$\begin{aligned} D(a_1, a_2, ..., a_n) &= -D(a_2, a_1, ..., a_n) \\ D(a_1, a_2, ..., a_n) + D(a_1, a_2, ..., a_n) &= 0 \end{aligned}$$

5.6 Exercise 6

We can swap rows and columns until A is in the same form as in Lemma 2. Since each row and column swap is equivalent to applying a transposition, we get that $\det A = (-1)^{i+j} \det A_{ij}$.

5.7 Exercise 7

Each term in the sum $D(a_1, ..., a_n) = \sum \sigma(p) a_{p_1 1} ... a_{p_n n}$ consists of exactly one element from each column and each row; swapping rows and columns does not change the terms in the sum. However, the permutation associated with each term is changed. The permutation p that sends $1 \to p_1$ becomes p' sending $p_1 \to 1$. This p' is exactly p^{-1}. Since $\sigma(1) = \sigma(p^{-1} \circ p) = \sigma(p^{-1})\sigma(p)$, $\sigma(p^{-1}) = \sigma(p)$ we are done.

5.8 Exercise 8

P is the linear transformation such that $P(e_j) = e_i$; in other words, P rearranges the representation of x by applying p to the components of x. We also have that $PQx = Pq(x) = p \circ q(x)$, since Qx permutes the components of x to produce $q(x)$, and $Pq(x)$ permutes the components of $q(x)$ to produce $p \circ q(x)$.

5.9 Exercise 9

$$\begin{aligned}
\operatorname{Tr} AB &= \sum_{i=1}^{m} (AB)_{ii} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ji} \\
\operatorname{Tr} BA &= \sum_{j=1}^{n} (BA)_{jj} = \sum_{j=1}^{n} \sum_{i=1}^{m} b_{ji} a_{ij}
\end{aligned}$$

5.10 Exercise 10

$$\begin{aligned}
\operatorname{Tr} AA^\top &= \sum (AA^\top)_{ii} = \sum \sum a_{ij} a^\top_{ji} \\
&= \sum \sum a_{ij}^2
\end{aligned}$$

6 Spectral Theory

6.1 Exercise 1

(a) We can re-express h as a linear combination of its eigenvectors, from which we can see that A^n causes all of these components to go to 0.

(b) Same as in part (a), except all of the components now go to ∞.

6.2 Exercise 2

$A(A^N f) = A(a^N f + Na^{N-1}h) = a^{N+1}f + a^N h + Na^N h$.

6.3 Exercise 3

Suppose $q(A) = \sum_{i=0}^{N} q_i A^i$. Then $q_i A^i f = q_i a^i f + q_i i a^{i-1} h$ by Exercise 2. From the linearity of the derivative it then follows that $q(A)f = q(a)f + q'(a)h$.

6.4 Exercise 4

Applying Lemma 9 to $p_1 ... p_k$ and p_{k+1} gives the desired result.

6.5 Exercise 5

$(A - aI)^d x = 0 \implies A(A - aI)^d x = 0 \implies (A - aI)^d (Ax) = 0 \implies Ax \in N_d$.

6.6 Exercise 6

Since m_A divides the characteristic polynomial of A (by definition of d_i), $m_A(A) = 0$ by Cayley-Hamilton. Suppose there is some polynomial $q(A) = 0$ with $\deg(q) < \deg(m_A)$. The roots of q must contain all of the eigenvalues of A, since for any eigenvector h we have that $q(A)h = q(a_h)h$ (where a_h denotes the eigenvalue associated with h). Thus, the roots of q can only differ in multiplicity from m_A. However, if any root of q has multiplicity $d_i' < d_i$, then we can choose an element $x \in N_{d_i}$ such that $q(A)x \neq 0$, which is a contradiction (since $N_{d_i'} \subset N_{d_i}$).

6.7 Exercise 7

The columns of A are Ax_i.

6.8 Exercise 8

By induction.

6.9 Exercise 9

The minimal polynomial of A divides the minimal polynomial of $A^\top$, and vice versa, so they must be the same. Thus, the indices of each eigenvalue of A and $A^\top$ must be the same, which means we can apply Theorem 12 to see that they are similar.

6.10 Exercise 10

$(\xi^{(i)}, x) = \sum k_j(\xi^{(i)}, x^{(j)}) = k_i(\xi^{(i)}, x^{(i)})$ since $(\xi^{(i)}, x^{(j)}) = 0$ by Theorem 17.