1 Sheldon Axler's Notes

Contents

2 General Notes

1. Recalling theorems shouldn't rely on photographic memory of equations, but a semantic recall that requires encoding the the theorem into their semantic components. 2. Being specific and creating your own names for the semantic relationship helps.

3 Chapter 1

3.1 $R^n and C^n$

1.1 Def: Complex Numbers
multiplication: $(a + bi)(c + di) = (ac - bd)(ad + bc)i$

3.2 Subspaces 1.C pg. 18

Subspaces
Definition: A subset that is a vector space.
Closed under operations of the vector space it subsides in.
1.40 Direct Sums
Sum of subspaces, such that each new vector sum must be a unique sum of the subspace's vectors.

$$U_1 \oplus \ ... \ \oplus \ U_n$$

3.2.1 Theorem 1.44 Condition for Direct Sum (Unproven)

Direct sum if and only if, the sum of an element from each subspace equaling 0, has only one option which is that each element is the 0 element. *proof* there is only one unique way to select a set of elements from each subspace s.t. their sum is 0, which is having each element to be hte zero vector. I don't know how this ties back into saying that the subspaces are then a direct sum for any other sum other than the 0 element.

4 Chapter 2

4.1 Basis 2.B p. 39

Definition: A set of linearly independent vectors of V that span V

2.29 Criterion of Basis: $v_1, ..., v_n$ are the basis of vector space V, if and only if v can be uniquely represented as:

$$v = a_1 v_1 + ... + a_n v_n$$
$$s.t. \ a_j \ \epsilon \ \mathbb{F}$$

Proof: (Forward) Assuming v_j from j = 1 to n is the basis, then v can be represented as a linear combination of it. Assume there are two possible limear combinations of scalars a's and c's. Subracting the linear combination results in the zero vector equaling a sum of $(a_j - c_j)v_j$. Since the vectors are linearly indepent, no vector can be a linear combination of the other, so all the individual differences of $a_j's$ and $c_j's$ will equal 0, so $a_j = c_j's$, so the both linear combinations are actually the same.
(Backward) Given all v's can be represented with the set of vectors $v_j's$, the span is implimitly proven. Since each linear combination for each v is unique, and since $v = 0 \epsilon V$, then these spanning vectors are linearly independent by defintion.

5 Chapter 3

5.1 Null Spaces and Range 3.B p.59

3.16 Linear map injective, iff null space 0.
3.20 Linear map surjective, iff range is codomain.
3.22 Fundamental Theorem of Linear Maps Given a finite dimensional vector space V:

$$\dim(\mathrm{V}) = \dim(\mathrm{range}(\mathrm{V})) + \dim(\mathrm{null}(\mathrm{V}))$$

Proof:
Dimension of V is cardinality of it's basis, of which the null space maps a subspace of vectors to 0, and the rest of the mapped vectors make up the range.

3.23 If dim(codomain) < dim(domain), then linear map cannot be injective, by definition of 3.16 and 3.20.

3.24 If dim(codomain) > dim(domain), then linear map cannot be surjective, by definition of 3.20 and 3.22.

3.26 Homogenous System

3.29 Inhomogenous System

5.2 Matrices 3.C p.70

3.30 Matrix
Definition: Matrix defined by (m, n) m = rows, n = columns. $M_{j,k}$ is the element at row j and column k.

3.32 Matrix as Linear Map
Matrix of linear map T, encodes $T(v_k)$ as a linear combination of W's basis vectors and column k of M.
Memory Tool: Draw domain basis at top of M, and codomain basis to the left of M.
For $T: F^n \to F^m$, assume k'th column of M(T) as T(k'th standard basis)

3.40 Dim($F^{m,n}$) = mn

5.3 Invertibility 3.D p.80

3.53 Invertibility
Definition: Given map S from V to W, there exists a map T from W to V, such that T(S) = I on V and S(T) = I on W.

3.54 Inverse is unique
Reasoning: There is only one unique way to flip the ordered mapping.
Proof: T has two inverses S 1 and 2 s.t. $S_1 = S_1 I = S_1(TS_2) = (S_1 T)S_2 = IS_2 = S_2$

3.55 Inverse of T is T^{-1}

3.56 Inverse if and only if Injective and Surjective.
Proof:
3.60 There is an isomorphism from $\mathcal{L}(V, W)$ and $F^{m,n}$

3.61 Dim($\mathcal{L}(V, W)$) = Dim(V)Dim(W)

3.58 Isomorphism
Invertible linear map between two vector spaces.
3.59 Dimension defines isomorphism
If two vector spaces have the same dimension, they are isomorphic.
Proof:

5.4 Product and Quotient Spaces 3.E p.91

3.71 Product Spaces Definition: V_1 x ... x V_n

$$(v_1, ..., v_n)\text{: } v_1 \epsilon V_1, ..., v_n \epsilon V_n$$

Addition on P.S.:$(v_1, ..., v_n) + (u_1, ..., u_n) = (v_1 + u_1, ..., v_n + u_n)$
Scalar multiplication on P.S.: distribute scalar λ to all elements

3.72 Product spaces are vector spaces
3.76 dim(product space) = sum(individual vector spaces)
Proof: Basis vector of V_j in j'th spot of product space with 0 else where, such that it's linearly idependent of all other basis vectors of other vector spacces.

3.79 Definition: $v + U$
Sum of vector and subspace v + U
$v + u :$ $u \epsilon U$ s.t. $v \epsilon V$ and U is subspace of V.

3.81 Def: affine subset, parallel
affine: $v + U$ is a affine subset of V
parallel: $v + U$ is parallel to U

3.83 Def: Quotient Space
set of all affine subjects of subspace U of V.
$V/U = v + U : v\epsilon V$

3.85 Theorem: two affine subsets either equal or disjoint
Given U of V and $v, w\epsilon of W$:
(a) $v - w\epsilon U$
difference of two elements of V are an element of the subset U. (b) $v+U = w+U$ the afffine subsets of U with respect to those two vectors are equal where each element is the same
(c) $(v + U) \cap (w + U) \neq$
the intersection of these two subsets is not the null set
proof $a \to b$ is shown by rearranging $v + u$ into the form of an element from $w+U$, where $v = w+(v-w)$ s.t. v-w is in the form of u so $v+u = w+((v-w)+u)$.

3.86 Def: + and scalar x on quotient space
addition: $(v + U) + (w + U) = (v + w) + U$
scalar mult: $\lambda(v + U) = (\lambda v) + U$

3.87 Theorem: Quotient space is Vector space
proof prove that addition and scalar multiplication of affine subsets are part of the quotient space, while also accounting for multiplicity of affine subsets, using 3.85 theorem on equivalence relations.

3.88 Def: Quotient Map
linear map $\pi : V \mapsto V/U$ s.t. $\pi(v) : v \mapsto v + U$ for $v\epsilon V$ s.t. U is a subspace of V

3.89 Theorem: Dimension of Quotient Space:
$dim(V/U) = dim(V) - dim(U)$
proof Defines range and null space of quotient map, then uses fundamental theorem of linear maps

3.90 Def: $\tilde{T}$ Linear Map
Given $T\epsilon\mathcal{L}(V, W)$, $\tilde{T} : V/null(T) \mapsto W$ s.t. $\tilde{T}(v + null(T) = T(v)$ *proof* make sure that the mapping is in fact linear, s.t. two elements of the same value map to the same output?

3.91 Theorem: Null space and Range of $\tilde{T}$ (a) $\tilde{T} : V/null(T) \mapsto W$ s.t. $T\epsilon\mathcal{L}(V, W)$
(b) $\tilde{T}$ is injective
(c) range(T)=range(T) (d) $V/null(T)$ is isomorphic to range(T) *proof* uses 3.85 to show that null space is zero, then by defintion c is true, and d is true by

implication of b and c.

5.5 Duality 3.F p. 101

3.92 Linear Functionals
Any linear transformation onto a field $\mathbb{F}$ that is an element from $\mathcal{L}(V, \mathbb{F})$

5.5.1 Def: 3.94 Dual Space V'

V' = $\mathcal{L}(V, \mathbb{F})$

Def. 3.95
Dim(V) = Dim(V')
Def. 3.96 Dual Basis of $v_1, ..., vn$ is set of functionals $\phi_1, ..., \phi_n$ s.t.

$$\phi_m(v_j) \ = \ 1 \text{ if } m \ = \ j \text{ or } 0 \text{ if } m \ \neq j$$

Def. 3.98 Dual basis of basis of V, is basis of V'
Proof Dual basis must span V' and be linearly independent.
Assume there exists a ϵ A, s.t.

$$\Phi \ = \ a_1\phi_1 + ... + a_n\phi_n \ = \ 0$$

Therefore, $\Phi(v_j) \ = \ a_j$. [How to prove all a's = 0] ?

6 Chapter 4

4.2 Definition
Given z = a + bi, real part Re(z) = a, and imaginary part Im(z) = b
Therefore, z = Re(z) + Im(z)i

4.3 Definition
Complex Conjugate: $\overline{z}$ = Re(z) - Im(z)i
Absolute value: $|z| = \sqrt{Re(z)^2 + Im(z)^2}$
*note: if z is real, z = $|z|$

4.5 Properties of Conjugates
Given $w, z\epsilon\mathbf{C}$,
i. sum of z and $|z|$: 2Re(z)
ii. diferrence of z and conjugate: 2Im(z)
iii. product of z and conjugate: $|z|^2$ iv. additivity and multiplicativity: conjugate of sum and product = product and sum of conjugates
v. conjugate of conjugate: cancels out
vi. imaginary and real parts bounded by $|z|$: by defintion

vii. absolute value of conjugate: equal to absolute value of original $|\overline{z}| = |z|$
viii. multiplicativity of absolute value: $|wz| = |w||z|$ s.t. square of a product equals product of squares
ix. triangle inequality: $|w + z| <= |w| + |z|$ s.t. proof requires taking square of left size of term and using property iii to prove inequality to the square of the right side of the term

4.6 Definition:
$p : \mathbf{F} \to \mathbf{F}$ with coefficients $a_0, ..., an\epsilon\mathbf{F}$ s.t. $p(z) = a_0 + a_1 z + a_2 z^2 + ... + a_n z^n \quad \forall z \epsilon \mathbf{F}$

4.7 Definition:
Zero polynomial is when all coefficients are 0.
Proof: Prove contrapositive, s.t. z >= 1, where it is defined as the sum of the absolute values of the coefficients over the absolute value of the last coefficient. Then by the definition of z, we prove that the last term is always greater than or equal to the n-1 terms of the polynomial, such that the function will always be some positive value.
*note: polynomials with two different sets of coefficients are unique, because subtracting the two would not be the zero function
*note: deg(0 polynimal) = $-\inf$, useful for later addition and multiplication operations

4.8 Definition:
Division Algorithm of Polynomials
Given $p, s\ \epsilon\mathcal{P}(\mathbf{F})$, s.t. $s \neq 0$, there exists a unique quotient and remainder, $q, r\epsilon\mathcal{P}(\mathbf{F})$, s.t. deg(r)<deg(s)
Proof : Uses statements from previous defitions of chapter three, to show that the polynomial's of degree p is a linear map from the product space of polynomials of the degree of the remainder (m-1) and divisor (n-m), for cases when the polynomial is greater than or equal to the divisor, $n \geq m$. Then prove it's an injection by showing the null space of the linear map is 0, then prove it is surjective showing that the range of the linear map is the same dimension as the dim of the polynomial p.

4.9 Definition:
Zero (or root) of polynomial, is $\lambda\epsilon\mathbf{F}$ s.t. for $p\epsilon\mathcal{P}(\mathbf{F})$ where $p(\lambda) = 0$.

4.10 Definition:
Factor of polynomial p is s, s.t. sq=p.

4.11 Definition:
Zero of a polynomial exits, if and only if, the zero is a degree 1 factor of the polynomial.
S.t. p(z) = $(z - \lambda)$q(z) for all z in P.
proof reverse direction is true, by plugging λ for z. forward direction is true

using division algorithm of polynomials 4.8.

4.12 Defintion:
Given a polynomial with degree $m \geq 0$, it has at most m zeros over it's field **F**
Proof: Using induction, using 4.11's definition, and recursing on the q polynomial which has a degree of m-1.
Base Case: prove cases for degree 0 and 1
Inductive Hypothesis: for $m > 1$, assume that each polynomial with degree m-1 has m-1 zeros
Inductive Step:

pre-note: the FTA proves existence.
4.13 Theorem:
Fundamental Theorem of Alebra
Every nonconstant polynomial with complex coefficients has at least one zero.
proof: uses Liouville's theorem to show the contrapositive of having no zeros, leads to a contradiction that the polynomial is a constant function.
note: polynomials of degree ≥ 5 don't have an equation, but there are numerical approximation techniques

4.14 Definition:
Factorization of Polynomial over **C**
given $p(z)\epsilon\mathcal{P}(\mathbf{C})$, there exists a unique $c, \lambda_1, ..., \lambda_n \epsilon \mathbf{C}$, s.t. $p(z) = c(z-\lambda_1)...(z-\lambda_n)$
proof used induction to prove existance and uniqueness. **gotlost!**

4.15 Defintion: Pairs of Zeros
Given a polynomial over complex field with real coefficients, it's zeros come in pairs
abbrev: p over C, w/ only R coef., then 0's in pairs
proof Used complex conjugate properites, s.t. the Re(complex #) is real.

4.15 Definition: Factorization of Quadratic
proof: pending!

4.16 Defintion: Factorization of Polynomial over **R**
proof: pending!

7 Chapter 5

7.1 5.A Invariant Subspaces

5.2 Defintion: Invariant Subspace
subspace U of V getting mapped by T, s.t. $\forall u \epsilon U, T(u) \epsilon U$

note: $T|_U$ is linear map T restricted to domain of subspace U
note: given $T\epsilon\mathcal{L}(V)$, 0, V, null(T), range(T) are subspaces – in particular, if element in domain is an element also found in the range, then it's invariant bc all elements from the domain get mapped to the range.

5.5 Defintion: eigenvalue
Given operator T over V, eigenvalue λ of T is value s.t. for $v\epsilon V$ & $v \neq 0$ where T(v)=λv
note: smallest invariant subspace U of operator T, s.t. dim(U)=1, is where each element u maps onto span(u), ergo eigenvector

5.6 Theorem: Equivalent Eigenvalue Definitions:
Given 5.5 that (a) λ is eigenvalue of T, then
(b) $T - \lambda I$ is not injective, surjective, or invertible
Proof: Since T is operator, 3.69 holds that if T is injective, then also surjective and invertible, and vice versa.

5.7 Definition: eigenvector
Suppose λ is eigenvalue of $T\epsilon\mathcal{L}(V)$. Then eigenvector of λ is all $v\epsilon V$ s.t. $v \neq 0$ and $T(v) = \lambda v$.
note: by defintion, eigenvector must be of $null(T - \lambda I)$

5.10 Theorem: Linearly Independent Eigenvectors
Given $T\epsilon\mathcal{L}(V)$, a set of m eigenvalues λ_i's set of corresponding eigenvectors are linearly independent.
proof Proof by Contradiction: assume m vectors are linearly dependent, then find first vector (indexed k) that is a linear combination of the previous vectors, then work with list of the k-1 vectors equaling zero with coefficients equaling to 0, then take difference of equation multiplieed by kth eigenvector and equaltion applied with linear transform T. This results in showing that all coefficients of a in the new equation equal zero because the k-1 eigenvalues are different from the kth eigenvalue, and this implies the linear combination of the k'th eigenvector equals 0 because all the a's are zero, and an eigenvector can't be zero ergo contradiction.

5.13 Theorem (5.12's Corollary): Maximum number of eigenvalues
Assuming $T\epsilon\mathcal{L}(V)$ s.t V is finite dimensional, then dim(V) is max. number of eigenvalues of T.
proof by 5.10, since eigenvectors are linearly independent they can be at most vectors as dimension of the space.

5.14 Def: Quotient and Restriction Operators
suppose: $T\epsilon\mathcal{V}$ and U is invariant subspace under V.
quotient operation: $T/U\epsilon\mathcal{L}(T/U)$ s.t. $(T/U)(v + U) = Tv + U$ for $v\epsilon V$
restriction operator: $T|_U\epsilon\mathcal{L}(U)$ s.t. $T|_U(u) = T(u)$ for $u\epsilon U$
proof restriction operator maps all U's back onto itself by the given assumption

so it's true; need to check for the multiplicity property of an affine subset to prove that it's still a vector space w̃hich I don't know how to do.
note: these definitions will allow for proofs later on.

7.2 5.B Eigenvectors of Upper-Triangular Matrix

5.16 Defintion: Powers of Operators
$T^m = T \cdots T$ m times
$T^0 = I$
$T^{-m} = (T^{-1})^m$

5.17 Defintion: Polynomial of Operators
Same definition of a polynomial over field **F** as in 4.6, but applied to operators instead of element of a field.
note: products of linear maps are composition functions
note: linear map $\mathcal{P}(\mathbf{F})$ to $\mathcal{L}(V)$ is defined as $p \mapsto p(T)$.

5.19 Definiton: Multiplication of Polynomials of **R**
(qp)(z)=q(z)p(z)

5.20 Theorem: Mult. Properties of Polynomials of Operators
a. multiplicativity
b. commutivity
proof **low on time!**

7.2.1 Theorem: 5.21 Existence of Eigenvalue on Operators of Complex Vector Space (UP)

Every operator on a (1) finite-dimensional, (2) non-zero, and (3) complex vector space has an eigenvalue.
proof Used Fundamental Theoream of Algebra to factorize a set of n+1 vectors that are are a a single vector v applied to the power of the operator. (low on time)

7.2.2 Def: 5.22 Matrix of Operator $\mathcal{T}(T)$

Given operator T over V, s.t $v_1, ..., v_n$ is a basis of V, then $M(T)_{j,k}$ (row, column) is defined by this linear combination relationship
$T(v_i) = A_{1,i}v_1 + ... + A_{n,i}v_n$ s.t. the linear combination of the basis of V is multiplied by the coefficients of the elements in a column.
note: goal is to simply matrix representation of a operator to be as simple as possible, "maximize the number of zeros"

7.2.3 Def: 5.24 Diagonal of Matrix

Given square matrix, it's all the entries of the matrix from top left to bottom right.

7.2.4 Def: 5.25 Upper-right Triangular Matrix

Square matrix where all the values below the diagonal are zero.

7.2.5 Theorem: 5.26 Invariant Subspaces of Upper Right Triangle Matrices

Given any condition, the others are true.
(a) There is an upper triangle matrix for $T \epsilon \mathcal{L}(V)$ w.r.t it's basis vectors $v_1, ..., v_n$ of V
(b) $T(v_j) \epsilon span(v_1, ..., v_j)$
(c) $span(v_1, ..., v_j)$ is an invariant subspace of V for each j

7.3 Eigenspaces and Diagonal Matrices 5.C

7.3.1 5.34 Def: Diagonal Matrix

matrix where all values except possibly diagonals are 0.

7.3.2 5.36 Def: Eigenspace

Given $T \epsilon \mathcal{L}(V)$ and $\lambda \epsilon \mathbb{F}$ then
eigenspace $E(\lambda, T) = null(T - \lambda I)$
therefore set of all eigenvectors of T along with zero vector.

note: does not require the field value to be an eigenvector & if you restrict the T to it's E, then it only performs multiplication by that field value & the eigenspace is written as a span of bectors

7.3.3 Theorem: 5.38 Sum of eigenspace is direct sum (UP)

given m distinct eigenvalues, the sum of it's corresponding eigenspaces on T is a direct sum
& and the sum of the dimension of each eigenspace is less than or equal to the dimensio of vector space V over which T is an operator
proof logical steps: (1) why do we suppose the sum of j different vectors from different eigenspaces equal zero (2) I understand how since they are eigenvectors from distinct eigenvalues that they are independent. ()

7.3.4 Def: 5.39 Diagonalizable

If an operator T has a diagonal matrix with respect to the basis of the vector space V it's mapped on, then it's diagonalizable.

8 Chapter 7: Operators on Inner Product Spaces

8.1 Positive Operators and Isometries 7.C

8.1.1 Def. 7.37: Isometry

Given operator $S \epsilon \mathcal{L}(V)$, if
$\|Sv\| = \|v\|$ a.k.a. preserves norm for every vector mapped

9 Xianglong's Operator Notes

1.6 Definiton: Minimal Polynomial
μ_A is min. poly. of A, if $\mu_A(A) = 0$ where A can be a square matrix or operator

1.8 Theorem: Division by Min. Poly.
p(A)=0 if.f then it is a multiple of the min. poly.

9.1 Determinant P.3

9.1.1 Axioms of Determinant 3.1

Note: Determinant is effect on volume 3 Axioms of Determinant of Matrix:
Given $F^{n,m} \to F$
1. $det(I_n) = 1$
2. Multilinearity
3. Alternating

9.2 Chage of Basis, Similar Matrices P.4

Note: $\mathcal{M}_\alpha$ means matrix with respect to basis α
Multiple basis representations for a matrix with same determinant, requires to prove the function is well-defined.
Assume $det(\mathcal{M}_\alpha) = det(\mathcal{M}_\beta)$
Understand composite function:
$\alpha \to \beta \to \beta \to \alpha$ s.t. switching of bases
Entire map is $\mathcal{M}_\alpha$ and middle map is $\mathcal{M}_\beta$

9.2.1 Definition 4.1: Change of Matrix Basis Form

Given each vector of basis α can be written as a linear combination of the basis of β.
Coefficients of linear combination of vector j from basis α is the column of field values in the matrix.
Note: $\mathcal{P}_{\alpha\beta}$ maps basis α to basis β

9.2.2 Proposition 4.2: Composite represented as Multiplication of Operators

9.2.3 Example 4.3: Change of Basis of (1,3),(1,4) to e_1, e_2

Done following corresponding matrixes of composite from 4.1, then applying multiplication from 4.2

9.2.4 Def 4.4: Similar Matrices