

MATH 110 Notes

based on *Linear Algebra Done Right*, third edition

Cheng, Feng

1 Vector Spaces

- A vector space V is an abelian group under addition that has the distributive and associative properties in scalar multiplication and the multiplicative identity. (Here addition and scalar multiplication are closed operations in V .)
- \mathbb{F}^S denotes the set of functions from S to \mathbb{F} . It is a vector space over \mathbb{F} under the sum and scalar multiplication of functions defined in the usual pointwise way.
- A vector space holds the cancellation law and has a unique additive identity and a unique additive inverse because it is a group. $0v = 0$, $a0 = 0$, and $(-1)v = -v$ hold by the distributive property.
- A subset U of the whole vector space V that is still a vector space is called a vector subspace. Usually we use the three criteria i) $0_V \in U$ (or U is nonempty), ii) closed under addition in U , and iii) closed under scalar multiplication in U to verify. To show the criteria are equivalent to the definition, in one direction we show that 0_V is the (unique) identity element of U and in the other direction we only need to check the additive inverse property (satisfied by $(-1)u = -u$).
- The sum $U_1 + U_2 + \cdots + U_m$, where U_i 's are subsets of V , is the set of all possible sums of the respective elements from each U_i .
- When the U_i 's are subspaces of V , then the sum is the smallest vector space containing all U_i 's.
- When the U_i 's are subspaces of V , and every element of the sum $U_1 + U_2 + \cdots + U_m$ can be written in only one way as $u_1 + u_2 + \cdots + u_m$, we call the sum a *direct sum*, denoted by \oplus replacing $+$.
- The equivalent criterion for direct sum is that the only way to express 0 is to write it as the sum of 0 's from each U_i . One direction is trivial, while the other direction employs the common trick of subtracting two different expressions of the same vector $v \in V$.
- If we are only discussing two subspaces U and W of V , then $U + W$ is a direct sum iff their intersection is the $\{0\}$. The proof of this, in both directions, considers $0 = v + (-v)$. The fact W always contains an element and its inverse will give us the answer.
- Arbitrary intersection of subspaces of V is a subspace of V , and the union of two subspaces is still a subspace iff one is contained in the other.

2 Finite-Dimensional Vector Spaces

Remark. Note that this section in the book builds on the notion of a list of vectors. FDVS is defined in terms of spanning list (which are the same), and the concept of a basis and the dimension of a vector space

is pushed backward. In our exposition, we rearrange the progression of these ideas and reconstruct some proofs.

- The *span* of a set S of vectors in V is defined to be all possible linear combinations of these vectors. In particular, $\text{span}(\emptyset) := \{0\}$.
- Similar to the fact that the sum $U_1 + U_2 + \cdots + U_m$ is the smallest vector space containing all U_i 's, $\text{span}(S)$ is the smallest subspace of V containing all the vectors in S . The proof of the two are almost exactly the same.
- A polynomial p in an indeterminate z with coefficients in \mathbb{F} is the **expression** $a_0 + a_1z + \cdots + a_nz^n$ with all a_i 's in \mathbb{F} . The *degree* of a polynomial is the largest exponent of z with nonzero coefficient, and if all the coefficients are 0, we customarily define its degree to be $-\infty$ in the book, or -1 elsewhere.
 - Since we are dealing with z and coefficients in $\mathbb{F} = \mathbb{R}$ or \mathbb{C} in this book, a polynomial here can be alternatively defined as the **function** $p : \mathbb{F} \rightarrow \mathbb{F}$ with $p(z) = a_0 + a_1z + \cdots + a_nz^n$. The reason behind is that a polynomial function with coefficients from an infinite field can be written into only one polynomial expression with its unique coefficients. See Chapter 4 of this book or Appendix E, Theorem 10 of *Linear Algebra* by Friedberg, Insel, and Spence. It is the consequence of the fact that a polynomial of degree $n \geq 1$ has at most n distinct zeros. We will proceed with this definition of polynomials.
 - We denote the set of all polynomials with coefficients in \mathbb{F} by $\mathcal{P}(\mathbb{F})$. Under the usual coefficient-wise addition and scalar multiplication, it is obvious that $\mathcal{P}(\mathbb{F})$ is a vector space over \mathbb{F} . In addition, the set of all polynomials with degree at most m , denoted by $\mathcal{P}_m(\mathbb{F})$ is a vector subspace of $\mathcal{P}(\mathbb{F})$.
- Linear independence and dependence can be characterized in multiple ways. A set S of vectors is linearly independent iff the only way to write 0 as $a_1v_1 + a_2v_2 + \cdots + a_mv_m$ is the trivial way. It is equivalent to saying that every linear combination has its only representation, or none of the vectors is in the span of the other vectors (can be written as a linear combination of them). Linear dependence is the negation of independence, but in particular we should remember that it means there exists at least one vector v_i that is in the span of the rest of the vectors.
 - Obviously removing such vector v_i from S will not change the span.
 - In particular, for a linearly independent set $S = \{s_1, s_2, \dots, s_m\} \subsetneq$ a linearly dependent set $T = \{s_1, s_2, \dots, s_m, t_1, t_2, \dots, t_n\}$, one such vector is always in $T \setminus S$. This is because for the nontrivial way of writing $0 = a_1s_1 + \cdots + a_ms_m + b_1t_1 + \cdots + b_nt_n$, if all b 's are 0, then all a 's will also be 0. Thus, some b_i must be nonzero, and the corresponding t_i can be removed without changing the span.
 - * The special case $n = 1$ of the contraposition will be useful to us. If one new vector is added to the linearly independent S but does not belong to $\text{span}(S)$, then the resulting set is still linearly independent.

- * From this special case we know that for a linear dependent set $\{v_1, \cdot, v_n\}$, at some index $k \leq n$ we have $\{v_1, \cdot, v_{k-1}\}$ being linearly independent, while $v_k \in \text{span}(\{v_1, \cdot, v_{k-1}\})$. (This is the well-ordering case of the linear dependence lemma 2.21 in book that we will use again in Chapter 5.)
- Note that a linearly independent set of vectors cannot have the 0 vector.
- We now have the essential tools to prove the following important theorem, that if a vector space V can be spanned by $S = \{w_1, w_2, \dots, w_n\}$, then all linearly independent sets of vector $\{u_1, u_2, \dots, u_m\} \subseteq V$ must have $m \leq n$. The proof suggests adding a new u_i into S and then remove another w_j from S in each step. In the end all the u 's will be added and the corresponding number of w 's are removed from S . Therefore, $m \leq n$.
 - The direct consequence of this theorem is the definition of dimension. But before that we need to introduce the notion of a basis and its existence.
- A *basis* of V is a linearly independent spanning set of V . A set is a basis iff all vectors in V can be written as a unique linear combination of the basis vectors. This follows directly from the definition.
- Most of time we are concerned with vector spaces with finite bases. The vector spaces that have a finite spanning set are called *finite-dimensional vector spaces* (FDVS), and we will soon know why they bear this name. First of all, we can show that every finite spanning set of a FDVS V can be reduced to a finite basis. Furthermore, every linearly independent set of a FDVS V can be extended to a basis.
 - To prove the first claim, we construct our basis B from scratch and try to add all the elements of the spanning set $S = \{v_1, v_2, \dots, v_n\}$ one by one into the basis, as long as the element added is not in the span of the existing vectors in B . In this way, every time a new vector is added to B , B will still be linearly independent, and this B will be linearly independent in the end. Furthermore, $\text{span}(B) = \text{span}(S) = V$ because all the v 's left out are in $\text{span}(B)$.
 - * It is also possible to start B from S and check from v_1 to v_n whether each vector is in the span of the rest of the vectors in B . If it is, then we remove it from B ; if it is not, then we keep it. In the end, $\text{span}(B) = \text{span}(S)$ and none of the vectors in B is in the span of the others. However, this is top-down approach is harder in practice to actually find the basis.
 - * This claim shows that every FDVS has a finite basis.
 - To prove the second claim, similarly construct our basis from the given linearly independent set L . We pick a finite basis $B = \{b_1, b_2, \dots, b_n\}$ of V and check whether the next b may be added to L (same to what we did above). In the end, all the b 's are in the span of L and L remains linearly independent.
- Choose two finite bases of a vector space. The fact that they are both linearly independent and both span V tells us that the two bases should have the same size, which we define as the *dimension* of a vector space. Thus, the vector spaces with finite bases are called *finite-dimensional*.
 - The fact that the bases of V has a fixed size implies every spanning set or linearly independent set of this size $\dim V$ is a basis itself (because the reduction/extension here is the trivial one).

- The subspace dimension $\dim U$ is always \leq the whole space dimension $\dim V$. To accomplish this we have to resort to the classic procedure: we add vectors from U that are not in the span of the existing ones in L . Thus L always remains linearly independent, and it should always be smaller in size than B , a spanning set of U . Our procedure will eventually terminate and no more vectors can be added to L , i.e., all the vectors in V is in $\text{span}(L)$. L is our desired basis, and $|L| = \dim U \leq \dim V = |B|$.
 - Note it is WRONG to assume that a finite basis of the whole space can always be reduced to a finite basis of the subspace!
 - Corollary: for FDVS V and its subspace U , if the two spaces have the same dimension, then they are the same space. This is an important result useful in many proofs.
- There are two more things that deserve mentioning. For FDVS V and its subspace U , we can always find subspace W such that $V = U \oplus W$. The idea is to find a basis B_U of U and extend it to the basis B_V of V . $\text{span}(B_V \setminus B_U)$ is our desired W . We then only have to show $V = U + W$ and $U \cap W = \{0\}$, which is not difficult.
- The dimension of the sum of two vector spaces U_1 and U_2 can be expressed as follows:

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

- Despite the fact that it looks like the inclusion-exclusion formula, $+$ and \cup are quite different, thus the formula cannot be similarly generalized to higher orders. To prove the formula, start from a basis $\{u_1, u_2, \dots, u_m\}$ of $U_1 \cap U_2$ (intersection is a subspace). U_1 and U_2 then have their respective bases B_1 and B_2 . We can show $B_1 \cup B_2$ is a basis of $U_1 + U_2$. See page 47 of the book for the full proof.
- Note that when the sum is a direct sum, then the formula is just $\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2$.

3 Linear Maps

3.1 The Vector Space of Linear Maps

- A linear map(ping)/transformation is a homomorphism of vector spaces under addition and scalar multiplication. The linear structure is preserved after the transformation T . To use symbols, a linear map is a function $T : V \rightarrow W$ such that for all $u, v \in V$ and $\lambda \in \mathbb{F}$ (the field of V and W),

$$T(u + v) = Tu + Tv \quad \text{and} \quad T(\lambda u) = \lambda T(u).$$

- The zero transformation maps every vector of V to 0_W . The identity transformation $I_V : V \rightarrow V$ maps each element of V to itself in V .
- We define the addition and scalar multiplication of linear maps in the vector-wise way:

$$(S + T)(v) = Sv + Tv \quad \text{and} \quad (\lambda T)(v) = \lambda T(v),$$

following which we can easily show that the set of linear maps $\mathcal{L}(V, W)$ from V to W is a vector space over \mathbb{F} . In addition to the linear properties, we define the *product* ST of linear maps same as $S \circ T$, given the appropriate domain and codomain. Why the product is essential will become clear later. The theory of functions tells us this product is associative $[T_1(T_2T_3) = (T_1T_2)T_3]$ and the identity works as usual ($I_W T = T I_V = T$). The distributive properties ($(S_1 + S_2)T = S_1T + S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$) obviously hold as well.

- If $T(x_1, \dots, x_n) = (a_{11}x_1 + \dots + a_{1n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n)$, then T is obviously a linear map from \mathbb{F}^n to \mathbb{F}^m . For a linear map $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$, it must be of this form, and all the a_{ij} 's are determined by the mapping of the n canonical basis vectors of \mathbb{F}^n .
- The uniqueness of the linear map that takes basis vectors v_1, v_2, \dots, v_n to corresponding vectors w_1, w_2, \dots, w_n is the most fundamental theorem in this section. It is our first step toward showing that linear maps and matrices are isomorphic. To prove the theorem, for arbitrary a_i 's, let

$$T : \sum_{i=1}^n a_i v_i \mapsto \sum_{i=1}^n a_i w_i.$$

- Here are a few properties that are useful in linear maps. The proofs are quite routine.
 - $T(0_V) = 0_W$;
 - T is linear iff $T(cx + y) = cT(x) + T(y)$ for all c, x, y .

3.2 Null Spaces and Ranges

- The *nullspace/kernel* $\text{null } T$ is the preimage of 0_W . The *range/image* $\text{range } T$ is defined the same as the range of a function. The nullspace and the range are respectively subspaces of V and W (by checking the three criterion). If the nullspace (resp. range) is finite-dimensional, then its dimension is the *nullity* (resp. *rank*) of T .
 - Since $\dim U = \dim W$ for subspace U of W implies $U = W$, $\text{rank } T = \dim W$ implies surjection.
 - A simple theorem to always keep in mind is that the mapping of the basis vectors of V spans $\text{range } T$.
- We are now endowed with all the tools to prove the **rank-nullity theorem**: for a FDVS V and T from V to W , $\text{range } T$ is finite-dimensional with

$$\dim V = \dim \text{null } T + \dim \text{range } T.$$

- As always, we need to construct a basis to work with first. Let u_1, u_2, \dots, u_n be a basis of $\text{null } T$, and thus can be extended to $u_1, \dots, u_m, v_1, \dots, v_n$ a basis of V . We prove $T(\{v_1, v_2, \dots, v_n\})$ is a basis for $\text{range } T$. The fact that the u 's are in the nullspace shows that Tv_1, \dots, Tv_n spans V . To show the elements are linearly independent, consider that

$$a_1 T(v_1) + \dots + a_n T(v_n) = 0$$

implies $a_1v_1 + \dots + a_nv_n \in \text{null } T$. Yet we also have $b_1u_1 + \dots + b_nu_n \in \text{null } T$, so we can set the two linear combinations equal. Because the u 's and v 's are altogether linearly independent, the coefficients a 's (and b 's) are all 0.

- The rank-nullity theorem gives rise to a number of equivalent characterizations regarding injection and surjection. First of all, there is a criterion that says T is injective iff $\text{null } T = 0$ (proof follows straight from the definition of nullspace). The results below follow.
 - A map to a smaller dimensional space cannot be injective because $\dim(\text{null } T) > 0$.
 - A map to a large dimensional space cannot be surjective because $\dim(\text{range } T) \leq \dim V < \dim W$.
 - V and W are FDVS of equal dimension, then

$$T \text{ is injective} \iff T \text{ is surjective} \iff \dim \text{range } T = \dim V.$$

- Solving systems of linear equations is a direct application of this.
 - A homogeneous system of linear equations with more variables than equations has nonzero solutions. Here we can view the system as a linear map T that maps the vector of unknown variables $x \in \mathbb{F}^n$ to $0 \in \mathbb{F}^m$ with $n > m$. The nullity of T is greater than 0 as a result.
 - Similarly, the T associated with an inhomogeneous equation with more equations than variables is not surjective, and thus some choices of constants on the right side will give no solutions to the system.

3.3 Matrices

- Consider the n -dimensional V with basis v_1, \dots, v_n and m -dimensional W with basis w_1, \dots, w_m . For $T \in \mathcal{L}(V, W)$, we can find for every v_k , a list of $A_{i,k}$'s such that $Tv_k = A_{1,k}w_1 + \dots + A_{m,k}w_m = \sum_{j=1}^m A_{j,k}w_j$. As a result, we have n lists of scalar a 's, each of length m , and we can put them into an $m \times n$ matrix A , with entries $a_{i,j}$ in the i -th row and j -th column. $\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$ uniquely determines the linear map $T \in \mathcal{L}(V, W)$.
 - Assume

$$\begin{aligned} Tv_1 &= A'_{1,1}w_1 + \dots + A'_{m,1}w_m, \\ &\vdots \\ Tv_n &= A'_{n,1}w_1 + \dots + A'_{n,m}w_m, \end{aligned}$$

from which we get the $m \times n$ matrix A' . Its transpose $n \times m$ A is our defined matrix representation of T . The reason for taking transpose will become clear later.

- We define matrix addition entry-wise, because under this definition, matrix representation of linear maps preserves addition, i.e. $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$ for $S, T \in \mathcal{L}(V, W)$, assuming $S + T$, S , and T share the same ordered bases.

- Similarly entry-wise scalar multiplication with matrices give us $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.
- Given the definition of matrix addition and scalar multiplication, $\mathbb{F}^{m,n}$, the set of matrices with m rows and n columns over \mathbb{F} , is a vector space of dimension mn .
- The reason why we represent linear maps from n -dimensional VS to m -dimensional VS is that we want to preserve the ordering of “ S ” and “ T ” after matrix representation, i.e., $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$ instead of $\mathcal{M}(T)\mathcal{M}(S)$, under the usual definition of matrix multiplication:

$$[AC]_{j,k} = \sum_{r=1}^n A_{j,r} C_{r,k}.$$

– Remember that matrix multiplication is not commutative, but is associative and distributive.

- The book defines $A_{j,\cdot}$ and $A_{\cdot,k}$ as the j -th row and the k -th column of matrix A , and thus $[AC]_{j,k} = A_{j,\cdot} C_{\cdot,k}$. Also, k -th column of matrix product equals matrix times the k -th column of the second matrix:

$$[AC]_{\cdot,k} = AC_{\cdot,k},$$

which is an alternative to understanding matrix multiplication. Obviously there is a row equivalent version that says j -th row of the matrix product equals j -th row times the second matrix.

- When we are considering matrix A times a column vector c , we can see Ac as the linear combination $c_1 A_{\cdot,1} + \cdots + c_n A_{\cdot,n}$.

3.4 Invertibility and Isomorphic Vector Spaces

- $T \in \mathcal{L}(V, W)$ is *invertible* if there exists S such that $ST = I_V$ and $TS = I_W$. We denote this inverse S by T^{-1} . One can show that the inverse must be linear and thus belongs to $\mathcal{L}(W, V)$. By the associativity of linear maps, the inverse is unique.
- Invertibility is equivalent to bijectivity between vector spaces. Linear maps are functions, so beyond showing that the inverse is linear, the proof is exactly the same.
- An *isomorphism* is an invertible linear map; if an isomorphism exists between two vector spaces, then the vector spaces are *isomorphic*.
- Two FDVS over \mathbb{F} are isomorphic iff they have the same dimension. Injection gives us nullity $= 0$ and surjection gives us rank $T = \dim W$. The rank-nullity theorem tells us $\dim V = \dim W$. On the other hand to prove isomorphism, we show $T(a_1 v_1 + \cdots + a_n v_n) = a_1 w_1 + \cdots + a_n w_n$ is both injective and surjective, which is not hard.

– The theorem above implies n -dimensional VS is always isomorphic to \mathbb{F}^n .

- Now we show the isomorphism between a linear map and its corresponding matrix. When the orders of the bases v_1, \dots, v_n and w_1, w_2, \dots, w_m are fixed, then \mathcal{M} is a function from $\mathcal{L}(V, W)$ to $\mathbb{F}^{m,n}$. As we have noted earlier \mathcal{M} is linear. Thus, it suffices to show that \mathcal{M} is bijective.

- For all $A \in \mathbb{F}^{m,n}$, let

$$Tv_k = \sum_{j=1}^m A_{j,k} w_j.$$

Then this is the *unique* (recall the theorem we emphasized in §3.1) T such that $A = \mathcal{M}(T)$.

- A linear map and its matrix representation are homomorphic not merely in regards to linearity, but also in regards to the preservation of the order of the product $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.
- By the theorems on isomorphism above, $\dim \mathcal{L}(V, W) = mn = (\dim V)(\dim W)$.
- We now shift our focus from the matrix representing an entire linear map to the matrix/column vector representing a vector in the same V . We fix the basis v_1, \dots, v_n , then for any $v \in V$, we define

$$\mathcal{M}(v) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix},$$

where a_1, \dots, a_n are given by $v = a_1 v_1 + \dots + a_n v_n$. Here \mathcal{M} is an isomorphism from V to $\mathbb{F}^{n,1}$.

- Following this definition, $\mathcal{M}(T)_{\cdot,k} = \mathcal{M}(Tv_k)$.
- More importantly, the linear map of a vector can now be completely described by matrix-multiplication. Symbolically, we want to prove $\mathcal{M}(Tv) = \mathcal{M}(T)\mathcal{M}(v)$ for all $v \in V$. Since $v = a_1 v_1 + \dots + a_n v_n$ and \mathcal{M} is linear,

$$\begin{aligned} \mathcal{M}(Tv) &= a_1 \mathcal{M}(Tv_1) + \dots + a_n \mathcal{M}(Tv_n) \\ &= a_1 \mathcal{M}(T)_{\cdot,1} + \dots + a_n \mathcal{M}(T)_{\cdot,n} \\ &= \mathcal{M}(T)\mathcal{M}(v). \end{aligned}$$

- How do we interpret this? Consider the following diagram under the appropriate bases for V and W .

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \mathcal{M} \uparrow \simeq & & \mathcal{M} \uparrow \simeq \\ \mathbb{F}^{n,1} & \xrightarrow{\mathcal{M}(T)} & \mathbb{F}^{m,1} \end{array}$$

For a vector $v \in V$, its image $\mathcal{M}(v)$ under the isomorphism \mathcal{M} of V belongs to $\mathbb{F}^{n,1}$. Matrix multiplication with $\mathcal{M}(T)$ gives us $\mathcal{M}(Tv) \in \mathbb{F}^{m,1}$, which we identify with Tv under the isomorphism \mathcal{M} of W . This is why a linear map can be described entirely by its matrix representation.

- If $V = \mathbb{F}^n$ and $W = \mathbb{F}^m$ and we are using the canonical bases, then the linear map is equivalent to matrix multiplication. Think about the Jacobian matrix representation of the total derivative in the Euclidean space. Left-multiplication by a matrix corresponds to the linear *transformation* of a vector.
- A *linear operator* is a linear map from a VS to itself. The notation $\mathcal{L}(V)$ stands for $\mathcal{L}(V, V)$.

- Suppose V is FDVS and $T \in \mathcal{L}(V)$, then (a) T is invertible iff (b) T is injective iff (c) T is surjective. (This is not true for infinite-dimensional vector space.) This is true for functions $f : S \rightarrow S$ (where S is finite) in general, but for linear maps the proof is likely to be rather simple:
 (a) \implies (b) we proved earlier.
 (b) \implies (c) by checking the rank-nullity theorem. Injection means nullity $T = 0$, which implies $\text{rank } T = \dim V$.
 (c) \implies (a) by showing T is injective, as nullity $T = \dim V - \text{rank } T$, which is 0.

3.5 Products and Quotients of Vector Spaces

- The *product* $V_1 \times \cdots \times V_n$ of vector spaces V_1, \dots, V_n over \mathbb{F} is the set

$$\{(v_1, \dots, v_n) \mid v_1 \in V_1, \dots, v_n \in V_n\}.$$

The addition and scalar multiplication on the product is defined in the usual entry-wise way. It is routine to prove that this product is a vector space over \mathbb{F} as well.

- $\dim(V_1 \times \cdots \times V_n) = \dim V_1 + \cdots + \dim V_n$. Why? Choose a basis for each V_i , then put every basis vector of V_i into the i -th entry of (v_1, \dots, v_n) and let other entries be 0. These vectors are linearly independent and span the product and thus is a basis of the product.
- For subspaces U_1, \dots, U_n of V , define the linear map $\Gamma : U_1 \times \cdots \times U_n \rightarrow U_1 + \cdots + U_n$ by

$$\Gamma(u_1, \dots, u_n) = u_1 + \cdots + u_n.$$

(One can easily check that this map is indeed linear.) Then $U_1 + \cdots + U_n$ is a direct sum iff Γ is injective (or invertible, since Γ is surjective by its definition.) To prove this, we just need to show null $T = 0$ iff the only way to write $0 = u_1 + \cdots + u_n$ is to let all u 's be 0. This is quite clear.

- The last two theorems lead to the following one. A sum of subspaces U_i 's of V is a direct sum iff dimensions add up. This is because two isomorphic vector spaces must have the same dimension. Therefore,

$$\dim(U_1 + \cdots + U_n) = \dim(U_1 \times \cdots \times U_n) = \dim U_1 + \cdots + \dim U_n.$$

Construct the linear map Γ above from the product of subspaces makes the proof really clean. Recall that we have proved the special case $n = 2$ already.

- An *affine subset* of V is a subset of V of the form $v + U = \{v + u \mid u \in U\}$ for some vector $v \in V$ and some subspace U of V . Here the affine subset $v + U$ is said to be *parallel* to U .
 – $v + U$ is a subspace of V iff $v \in U$. If $v \in U$, then $v + U = U$, subspace of V . If $v \notin U$, then $0 \notin v + U$, and thus $v + U$ cannot be a subspace of V .
- Under the same setting, the quotient space V/U is the set of all affine subsets of V parallel to U , i.e., $V/U = \{v + U \mid v \in V\}$.

- For subspace U of V and $v, w \in U$, then

$$(a) \ v - w \in U \iff (b) \ v + U = w + U \iff (c) \ (v + U) \cap (w + U) \neq \emptyset.$$

(a) \implies (b) as $v + u = w + ((v - w) + u) \in w + U$ (and the same for the other direction). (b) \implies (c) because $v + U$ is not empty, while (c) \implies (a) because $v + u_1 = w + u_2$ for some u_1 and u_2 , and we then move $v - w$ to the left side.

– The theorem above tells us that two affine subset to U are either the same or disjoint.

- The addition and scalar multiplication on V/U are defined by

$$(v + U) + (w + U) = (v + w) + U \quad \text{and} \quad \lambda(v + U) = \lambda v + U.$$

We should pay particular attention to this definition because **an affine subset do not have a unique representation** (for $v \neq v'$, it is possible that $v + U = v' + U$ still)! Therefore, one has to show that different representations $v + U$ and $w + U$ of the same affine subsets of U still gives us a **unique** $(v + w) + U$, so that the addition function above is well-defined (and similar for scalar multiplication). Suppose $v + U = v' + U$ and $w + U = w' + U$, then essentially we want to prove $(v + w) + U = (v' + w') + U$. By assumption $v - v'$ and $w - w' \in U$, $(v + w) - (v' + w') \in U$, and we reach the conclusion (and similar for scalar multiplication).

- V/U with the two operations above is a vector space. Showing this is routine, but one has to remember that V/U is not a subspace of V . Despite all the elements of the quotient space is in the whole space, the additive identity is now $0 + U = U$, and the additive inverse of $v + U$ is $-v + U$. (The quotient space is discussed with respect to a fixed subspace.) Therefore, we have to check the 8 VS properties.

– Also, by the previous theorem, for any $v \in U$, $v + U$ equals the 0 vector $(0 + U)$ and is thus not a basis vector for V/U .

- To find the dimension of V/U , we appeal to the trick of constructing a linear map again. (One should check its linearity) The *quotient map* $\pi : V \rightarrow V/U$ is given by

$$\pi(v) = v + U.$$

For FDVS V and its subspace U , $\dim V/U = \dim V - \dim U$. π is a map onto V/U obviously, and null $\pi = U$. Therefore the formula follows.

- Now we define the induced map $\tilde{T} : V/(\text{null } T) \rightarrow W$ of $T : V \rightarrow W$ by

$$\tilde{T}(v + \text{null } T) = Tv.$$

$u + \text{null } T = v + \text{null } T$ implies $u - v \in \text{null } T$, and thus $Tu - Tv = 0$. Therefore the induced map is

well defined. Checking the linearity of \tilde{T} is routine. Note that $T = \tilde{T} \circ \pi$.

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ & \searrow \pi \quad \nearrow \tilde{T} & \\ & V/\text{null } T & \end{array}$$

- The definition leads to the isomorphism theorem that exists for different algebraic structures. For $T \in \mathcal{L}(V, W)$, \tilde{T} is injective and $\text{range } \tilde{T} = \text{range } T$. The latter simply follows from the definition, while the first requires us to prove $\text{null } \tilde{T} = \{0 + \text{null } T\}$, the identity. $\tilde{T}(v + \text{null } T) = Tv = 0$ gives us $v - 0 \in \text{null } T$. Hence, $v + \text{null } T = 0 + \text{null } T$, and the conclusion follows.

– The two claims combined tell us $V/(\text{null } T) \simeq \text{range } T$ under the isomorphism \tilde{T} .

3.6 Duality

- A *linear functional* on V is a linear map belonging to $\mathcal{L}(V, \mathbb{F})$. The *dual space* of V , denoted usually by V^* or V' , is the vector space $\mathcal{L}(V, \mathbb{F})$.
- We proved earlier that $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$. Therefore, $\dim V' = \dim V$ for finite-dimensional V .
- The *dual basis* of v_1, \dots, v_n , basis of V , is a set of vectors $\varphi_1, \dots, \varphi_n$ in V' , where each φ_i is a linear functional defined by

$$\varphi_i(v_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

These φ_i 's are designed to form a basis of V' . It suffices to show that $\varphi_1, \dots, \varphi_n$ is linearly independent in V' . Consider

$$a_1\varphi_1 + \dots + a_n\varphi_n = 0.$$

Since $(a_1\varphi_1 + \dots + a_n\varphi_n)(v_j) = a_j$, all a_j 's are 0.

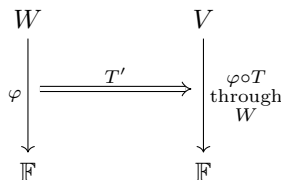
– Remember that the dual basis vector φ_i of e_1, \dots, e_n in \mathbb{F}^n has

$$\varphi_i(x_1, \dots, x_n) = \varphi_i(x_1e_1 + \dots + x_ne_n) = x_i.$$

Namely, the i -th dual basis vector maps (x_1, \dots, x_n) to its i -th coordinate.

- For v_1, \dots, v_n basis of V and its associated dual basis $\varphi_1, \dots, \varphi_n$ of V' ,
 - first, every $\psi \in V'$ can be written as $\sum_{i=1}^n \psi(v_i)\varphi_i$. Since $(\sum_{i=1}^n \psi(v_i)\varphi_i)(v_j) = \psi(v_j)$, the sum and ψ agrees on all basis vectors and they are equal.
 - second, every $v \in V$ can be written as $\sum_{i=1}^n \varphi_i(v)v_i$. For this claim one needs to prove that for $v = a_1v_1 + \dots + a_nv_n$, every $a_j = \varphi_j(v)$. By linearity, $\varphi_j(v) = \sum_{i=1}^n a_i\varphi_j(v_i) = a_j$.

- The *dual map* $T' \in \mathcal{L}(W', V')$ of $T \in \mathcal{L}(V, W)$ is given by $T'(\varphi) = \varphi \circ T$ for $\varphi \in W'$. We have to be very clear about what this definition means. Here T' maps a function to a function.



T' is a composition of linear maps and is thus a linear map from V' to \mathbb{F} . To show that T' is linear from W' to V' , one simply follows the definition.

– The following algebraic properties hold for dual maps:

- * $(S + T)' = S' + T'$,
- * $(\lambda T)' = \lambda T'$,
- * $(ST)' = T' S'$

under the appropriate S, T , and λ . The first two are routine, and the third one is manipulation of function compositions.

- We define *annihilator* U^0 of $U \subseteq V$ as $\{\varphi \in V' \mid \varphi(U) = 0\}$, the subset of all linear functionals on V whose image is 0.
- There is a good example (3.104) in the book about annihilators. For the dual basis $\varphi_1, \dots, \varphi_5$ in $(\mathbb{F}^5)'$ that corresponds to e_1, \dots, e_5 in \mathbb{F}^5 , $(\text{span}(e_1, e_2))^0 = \text{span}(\varphi_3, \varphi_4, \varphi_5)$, since φ_i takes out the i -th coordinate.
- As expected U^0 is a subspace of V' . The zero linear functional takes $U \subseteq V$ to 0. $(\varphi + \psi)(U) = 0$ and $\lambda\varphi(U) = 0$ are omitted as usual.
- For FDVS V and its subspace U ,

$$\dim U + \dim U^0 = \dim V.$$

The book mentions two proofs. The first proof takes the usual “choose a basis and extend it” approach. The second proof is more obscure: consider the inclusion map $i \in \mathcal{L}(U, V)$ such that $i(u) = u$ for all $u \in U$, then $i' \in \mathcal{L}(V', U')$, which gives us

$$\dim \text{null } i' + \dim \text{range } i' = \dim V' = \dim V.$$

Note that $i'(\varphi) = \varphi \circ i$ for $\varphi \in V'$, giving us $\text{null } i' = U^0$. The nullspace of i' are the set of φ 's such that $\varphi \circ i = 0$, and therefore φ should map the entire U to 0.

Now we show $U = \text{range } i'$. A linear map defined on a subspace can always be extended to the whole space, and thus we may extend any $\varphi \in U'$ to $\psi \in V'$. The dual map i' maps ψ back to φ , so $\varphi \in U' \implies \varphi \in \text{range } i'$, showing us what we desire.

- The nullspace and range of the dual map T' can now be characterized in terms of the concepts we have introduced.

- For $T \in \mathcal{L}(V, W)$, $\text{null } T' = (\text{range } T')^0$:

$$\varphi \in \text{null } T' \iff (\varphi \circ T)(v) = \varphi(Tv) = 0 \iff \varphi \in (\text{range } T)^0.$$

- If V, W are finite-dimensional, then we have

$$\begin{aligned} \dim \text{null } T' &= \dim(\text{range } T)^0 \\ &= \dim W - \dim \text{range } T \\ &= \dim W - (\dim V - \dim \text{null } T) \\ &= \dim \text{null } T + \dim W - \dim V. \end{aligned}$$

- $\dim \text{range } T' = \dim \text{range } T$ easily follows from the rank-nullity theorem and the two bullet points above.
- Also we can prove $\text{range } T' = (\text{null } T)^0$. Showing $\text{range } T' \subseteq (\text{null } T)^0$ is standard (just assume T' maps φ to $\psi \in \text{range } T'$), while the other direction is quite tricky. The book assumes finite-dimension here and showed the two spaces have the same dimension:

$$\dim \text{range } T' = \dim \text{range } T = \dim V - \dim \text{null } T = \dim(\text{null } T)^0.$$

We have to make a side note here. When we proved the first bullet point, we did not assume finite-dimension of V and W , so it is natural to consider if our proof of $\text{range } T' \supseteq (\text{null } T)^0$ also does not require the finite-dimension assumption either.

We may formulate our question this way: for $T \in \mathcal{L}(V, W)$ and $\varphi \in V'$ with $\varphi(\text{null } T) = 0$, can we always find a $\psi \in W'$ such that $\varphi = \psi \circ T$? This turns out to be true if we assume the axiom of choice from set theory, which has important corollaries in the theory of infinite-dimensional vector spaces. We will not get into details here, but it is interesting to know about this.

- The corollary of the previous claims are the following:

$$T \text{ is surjective} \iff T' \text{ is injective} \quad \text{and} \quad T \text{ is injective} \iff T' \text{ is surjective}.$$

Proof of this is quite standard, and we show the first one here only:

$$T \in \mathcal{L}(V, W) \text{ surjective} \iff \text{range } T = W \iff (\text{range } T)^0 = 0 \iff \text{null } T' = 0 \iff T' \text{ injective}.$$

- The *row rank* and *column rank* of matrix A are defined to be the dimension of the span of the rows of A in $\mathbb{F}^{1,n}$ and of the columns of A in $\mathbb{F}^{m,1}$, respectively. As we know from lower-division matrix algebra, they are the same and represent the number of pivots in a reduced row echelon form of a matrix. We now prove this fact with the tools we have.

- $\mathcal{M}(T') = (\mathcal{M}(T))^t$. Consider the dual basis ψ_j ($1 \leq j \leq m$) of W' and φ_j ($1 \leq k \leq n$) of V' . Then

$$T'(\psi_j) = \sum_{r=1}^n C_{r,j} \varphi_r,$$

which we evaluate at v_k :

$$(\psi_j \circ T)(v_k) = \sum_{r=1}^n C_{r,j} \varphi_r(v_k) = C_{k,j}.$$

On the other hand, we have

$$(\psi_j)(Tv_k) = \psi_j\left(\sum_{r=1}^m A_{r,k} w_r\right) = \sum_{r=1}^m A_{r,k} \psi_j(w_r) = A_{j,k}.$$

- First, $\dim \text{range } T = \text{column rank of } \mathcal{M}(T)$. From bases v_1, \dots, v_n and w_1, \dots, w_m of V and W , we have $\text{span}(\{Tv_1, \dots, Tv_n\})$ is isomorphic to $\text{span}(\{\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n)\})$ under the isomorphism \mathcal{M} . Thus, the two VS have the same dimensions, with $\dim \text{span}(\{\mathcal{M}(Tv_1), \dots, \mathcal{M}(Tv_n)\}) = \text{column rank of } \mathcal{M}(T)$. Recall that $\text{span}(\{Tv_1, \dots, Tv_n\}) = \text{range } T$, thus $\dim \text{range } T = \text{column rank of } \mathcal{M}(T)$, as desired.
- For $A \in \mathbb{F}^{m,n}$, define $T \in \mathcal{L}(V, W)$ as the left-multiplication by A map (i.e., $Tx = Ax$). Then $A = \mathcal{M}(T)$ with respect to the canonical bases. It follows that

$$\begin{aligned} \text{column rank of } A &= \text{column rank of } \mathcal{M}(T) \\ &= \dim \text{range } T = \dim \text{range } T' \text{ (which we proved a moment ago)} \\ &= \text{column rank of } \mathcal{M}(T') = \text{column rank of } A^t = \text{row rank of } A. \end{aligned}$$

Therefore, *rank* is a property of a finite matrix.

4 Polynomials

Remark. We used facts about polynomials in high school algebra. These facts were intuitively “understood” by us back then, but we never had the chance to rigorously prove them. Although the theory of polynomials is definitely something one should know at this stage, we will skim through this chapter in our notes. Please refer to Chapter 4 of our textbook or Appendix E of *Linear Algebra* by Friedberg, Insel, and Spence for the detailed development of major results.

- The division algorithm for polynomials is similar to the one in number theory. For $p, s \in \mathcal{P}(\mathbb{F})$ with $s \neq 0$, then there exist a pair of unique polynomials $q, r \in \mathcal{P}(\mathbb{F})$ such that $p = sq + r$ and $\deg r < \deg s$.
 - With the theory of linear maps this becomes easy to prove.
- Each zero of a polynomial corresponds to a degree-1 factor, i.e., for $p \in \mathcal{P}(\mathbb{F})$ and $\lambda \in \mathbb{F}$, $p(\lambda) = 0$ iff $\exists q \in \mathcal{P}(\mathbb{F})$ such that for all $z \in \mathbb{F}$, $p(z) = (z - \lambda)q(z)$. To prove the “ \implies ” direction, division

algorithm gives us $p(z) = (z - \lambda)q(z) + r$ (here r is a constant) as $(z - \lambda)$ is of degree 1. Thus, $p(\lambda) = 0$ gives us $r = 0$.

- A polynomial $p \in \mathcal{P}(\mathbb{F})$ has at most as many zeros as its degree (≥ 0). $\deg = 0$ and 1 can be checked, and we proceed with induction on higher orders. If p has no zeros, then done. If p has one zero, let

$$p(z) = (z - \lambda)q(z),$$

where p is of degree $m + 1$ and q is of degree m (with at most m zeros).

- Recall the remark on polynomials we made in Chapter 2. Now we prove that polynomial functions over an infinite field \mathbb{F} have a unique representation. Suppose $f, g \in \mathcal{P}(\mathbb{F})$ and $f(a) = g(a)$ for all $a \in \mathbb{F}$, we want to show that $f = g$.

Define $h = f - g$, and then we have $\forall a \in \mathbb{F}, h(a) = 0$. Suppose h is a nonzero polynomial of degree n , then by the division algorithm h has at most n zeroes, yet \mathbb{F} is an infinite field, contradiction! Thus $h = f - g$ is the zero polynomial, and $f = g$ as a result.

- The proof of the fundamental theorem of algebra is omitted here for the obvious reason. Every nonconstant polynomial on \mathbb{C} has a zero, which directly implies that a polynomial p of degree m has m nondistinct roots, which we sketch through now.

- Every nonconstant polynomial over \mathbb{C} has a unique factorization (up to the order of factors) of the form

$$p(z) = c(z - \lambda_1) \cdots (z - \lambda_m),$$

where $c, \lambda_1, \dots, \lambda_m \in \mathbb{C}$.

- To prove this, by the fundamental theorem of algebra we can factorize $p(z)$ into $(z - \lambda)q(z)$, and again factorize $q(z)$, and so on. We have proved the existence of such factorization.
- Now we proceed to prove the uniqueness.
- For polynomials $p \in \mathcal{P}(\mathbb{C})$ with real coefficients, the complex conjugate of a root is still a root. This follows from $\lambda^m = \bar{\lambda}^m$.
- For nonconstant $p \in \mathcal{P}(\mathbb{R})$, p has a unique factorization into linear and quadratic polynomials with the highest coefficients being 1, i.e.,

$$p(x) = c(x - \lambda_1) \cdots (x - \lambda_m)(x^2 + b_1x + c) \cdots (x^2 + b_Mx + c_M),$$

where c , the λ 's, the b 's, and the c 's are all reals and each quadratic factors have negative discriminant.

5 Eigenvalues, Eigenvectors, and Invariant Subspaces

5.1 Invariant Subspaces

- For $T \in \mathcal{L}(V)$, the subspace U of V is *T-invariant* if $T(U) \subseteq U$, or alternatively, $T|_U$ is an operator on U .
- It can be easily verified that $\{0\}$, V , $\text{null } T$, and $\text{range } T$ are invariant subspaces of V under T .
- We now focus on invariant subspaces of dimension 1. Any 1-dimensional subspace U has some vector $u \in V$ as the basis (actually any vector in U can be the basis). To check if T is invariant under U , we only need to check if the basis vectors are mapped into U (so that the range $T \subseteq U$). Since $Tu \in U$ iff $Tu = \lambda u$ for some $\lambda \in \mathbb{F}$, a subspace with dimension 1 is invariant iff a vector u in the subspace has $Tu = \lambda u$.
- For $T \in \mathcal{L}(V)$, $\lambda \in \mathbb{F}$ is an *eigenvalue* of T if there exists a nonzero $v \in V$ such that $Tv = \lambda v$. The v here is called the *eigenvector*.
- To move λv to the left side, we consider the linear operator λI that maps v to λv . Therefore, we have $(T - \lambda I)v = 0$. Since $v \neq 0$, $T - \lambda I$ is not injective/surjective/invertible (since it is an operator on V).
- Eigenvectors corresponding to distinct eigenvalues are linearly independent. The usual proof of this starts with 2 distinct eigenvalues and proceeds with induction, but the book uses the well-ordering principle instead. Suppose the v_1, \dots, v_n corresponding to eigenvalues $\lambda_1, \dots, \lambda_n$ are linearly dependent, then at $k \leq n$ we have $v_k \in \text{span}(v_1, \dots, v_{k-1})$ with v_1, \dots, v_{k-1} being linearly independent. Thus,

$$v_k = a_1 v_1 + \dots + a_{k-1} v_{k-1}, \quad (\star)$$

and by applying T to both sides, we have

$$\lambda_k v_k = a_1 \lambda_1 v_1 + \dots + a_{k-1} \lambda_{k-1} v_{k-1}.$$

If instead, we multiply (\star) by λ_k on both sides, then

$$\lambda_k v_k = a_1 \lambda_k v_1 + \dots + a_{k-1} \lambda_k v_{k-1}.$$

Subtracting the two equations above give us

$$0 = a_1(\lambda_1 - \lambda_k) + \dots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1}.$$

Since v_1, \dots, v_{k-1} are linearly independent, all the coefficients are 0 and thus all the a 's are the same (because the λ 's are distinct). However, this implies $v_k = 0$, which contradicts with v_k being an eigenvector. Therefore, the set of eigenvectors are linearly independent.

- The direct corollary of this is that the number of eigenvalues of $T \in \mathcal{L}(V)$ cannot exceed the $\dim V$.

- Here we restate the $T|_U \in \mathcal{L}(U)$ as the *restriction operator* of T . The *quotient operator* $T/U \in \mathcal{L}(V/U)$ is given by $(T/U)(v + U) = Tv + U$ for any $v \in V$. Of course for the quotient operator we need to verify it is well-defined for different representation of the same quotient space, which is very similar to what we have done before and is thus omitted. These two operators provide nice information about the original T .

5.2 Eigenvectors and Upper-Triangular Matrices

- Linear operators allow idempotent operations. Given $T \in \mathcal{L}(V)$, we define T^m ($m \in \mathbb{Z}^+$) as T composite with itself m times, T^0 as the identity operator I_V , and T^{-m} as $(T^{-1})^m$, given T^{-1} exists.
 - Obviously $T^m T^n = T^{m+n}$ and $(T^m)^n = T^{mn}$. (If T is invertible then we can m, n are allowed to be all integers.)
- $p(T)$ is the polynomial p taking T as its indeterminate (although T is not \mathbb{F}). For $p(z) = a_0 + a_1 z + \dots + a_n z^n$ over \mathbb{F} , we define $p(T) = a_0 I + a_1 T + \dots + a_n T^n$.
 - If we fix an operator T , the evaluation map $\mathcal{P}(\mathbb{F})$ to $\mathcal{L}(V)$ given by $p \mapsto p(T)$ is linear.
- The product of $pq \in \mathcal{P}(\mathbb{F})$ of polynomials $p, q \in \mathcal{P}(\mathbb{F})$ is defined by

$$(pq)(z) = p(z)q(z)$$

for all $z \in \mathbb{F}$. If we write out the polynomials, one can easily see that $(pq)(z) = p(z)q(z) = q(z)p(z) = (qp)(z)$.

- It follows directly that $(pq)(T) = p(T)q(T)$ and $(pq)(T) = (qp)(T)$, because we only replace the symbol z by T .
- It turns out that every operator on a finite-dimensional and nonzero complex vector space necessarily has an eigenvalue. (This is not necessarily true on real vector spaces) The proof employs the fundamental theorem of algebra and is quite tricky. Choose a nonzero $v \in V$ ($\dim V = n$) and consider the list $v, Tv, \dots, T^n v$, which must be linearly dependent. Thus, we have

$$0 = a_0 v + a_1 Tv + \dots + a_n T^n v,$$

where the a_i 's are not all 0. In particular, a_1, \dots, a_n cannot all be 0 since $v \neq 0$ will then give us $a_0 = 0$. Thus, we have a nonconstant polynomial $a_0 + a_1 z + \dots + a_m z^m$ ($m \leq n$), which can be factorized into

$$c(z - \lambda_1) \cdots (z - \lambda_m)$$

over \mathbb{C} . For

$$0 = c(T - \lambda_1 I) \cdots (T - \lambda_m I)v,$$

the nonzero vector $w = (T - \lambda_{j+1} I) \cdots (T - \lambda_m I)v$ will be taken to 0 by $(T - \lambda_j I)$. $(T - \lambda_j I)$ is thus not injective and λ_j is our eigenvalue.

- The matrix of an operator is a square matrix. A square matrix with all entries being 0 below the diagonal is called an *upper-triangular matrix*, which is crucial in the theory of eigenvalues.
- How are eigenvalues and vectors, invariant subspaces, and upper-triangular matrices connected? Suppose $T \in \mathcal{L}(V)$ and v_1, \dots, v_n is a basis of V , then

- (a) $\mathcal{M}(T)$ is upper triangular.
- (b) $Tv_j \in \text{span}(v_1, \dots, v_j)$ for each j from 1 to n .
- (c) $\text{span}(v_1, \dots, v_j)$ is invariant under T for each j from 1 to n .

The first two are equivalent follow straight from the definition. (b) \implies (c) because for every i between 1 and j , $Tv_i \in \text{span}(v_1, \dots, v_j)$. (c) \implies (b) is obvious as well.

- For FDVS V over \mathbb{C} and $T \in \mathcal{L}(V)$, then $\mathcal{M}(T)$ is upper triangular with respect to some basis of V .
- Upper-triangular matrices help us determine whether T is invertible. If $T \in \mathcal{L}(V)$ has an upper-triangular $\mathcal{M}(T)$ w.r.t. some basis of V , then T is invertible iff all entries on the diagonal of $\mathcal{M}(T)$ are nonzero.
- By the previous theorem, we can determine all the eigenvalues of T if we are fortunate to find a basis that corresponds to an upper-triangular $\mathcal{M}(T)$. $\mathcal{M}(T - \lambda I)$ is of the form

$$\begin{bmatrix} \lambda_1 - \lambda & & & * \\ & \lambda_2 - \lambda & & \\ & & \ddots & \\ 0 & & & \lambda_n - \lambda \end{bmatrix}.$$

One of the $\lambda_i - \lambda$'s is 0 iff $T - \lambda I$ is invertible iff λ is an eigenvalue. Therefore, all the entries λ_i 's on the diagonal are exactly all the eigenvalues.

5.3 Eigenspaces and Diagonal Matrices

- A *diagonal matrix* has all its entries off the diagonal 0.
- The *eigenspace* of $T \in \mathcal{L}(V)$ corresponding to $\lambda \in \mathbb{F}$ is given by

$$E(\lambda, T) = \text{null}(T - \lambda I) = \{v \in V \mid Tv = \lambda v\}.$$

The eigenspace is basically the vector subspace of V with all eigenvectors corresponding to λ , along with the 0 vector.

- The sum of eigenspaces is a direct sum. We consider $u_1 + \dots + u_n = 0$, where each eigenvector u_i is from the eigenspace $E(\lambda_i, T)$. Because eigenvectors corresponding to distinct eigenvalues are linearly independent, we get what we want. Furthermore, since

$$E(\lambda_1, T) + \dots + E(\lambda_m, T)$$

is a direct sum, we have

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_n, T) = \dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_n, T)) \leq \dim V.$$

- An operator T is *diagonalizable* if the operator has a diagonal matrix w.r.t. some basis of T .
- For T on FDVS V and distinct eigenvalues $\lambda_1, \dots, \lambda_m$, diagonalizability of T is equivalent to the following four conditions:
 - (a) V has a basis consisting of eigenvectors corresponding to each λ_i 's;
 - (b) V can be decomposed into n 1-dimension subspaces U_1 to U_n , where $V = U_1 \oplus \cdots \oplus U_n$;
 - (c) $V = E(\lambda_1, T) + \cdots + \dim E(\lambda_n, T)$;
 - (d) $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_n, T) = \dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_n, T))$

We consider the whole list of eigenvalues $\lambda_1, \dots, \lambda_n$ with repetition ($n \geq m$).

•

6 Inner Product Spaces

6.1 Inner Products and Norms

- For $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, the *dot product*

$$x \cdot y := x_1 y_1 + \cdots + x_n y_n,$$

which is a scalar in \mathbb{R}^n .

- The *inner product* is the abstraction of the dot product to more complicated products of vectors on V over \mathbb{R} and \mathbb{C} . It is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$ (\mathbb{R} or \mathbb{C}) with the following properties:
 - Positive definiteness: $\langle v, v \rangle \geq 0$, which achieves “=” iff $v = 0$;
 - Linearity in the first slot: $\langle \lambda u + v, w \rangle = \lambda \langle u, w \rangle + \langle v, w \rangle$;
 - Conjugate symmetry: $\langle u, v \rangle = \overline{\langle v, u \rangle}$.

The properties that directly ensue include:

- Conjugate symmetry gives us “conjugate linearity” in the second slot:

$$\langle u, \lambda v + w \rangle = \bar{\lambda} \langle u, v \rangle + \langle u, w \rangle;$$

- For fixed $w \in V$, $f : v \mapsto \langle v, w \rangle$ is a linear map from V to \mathbb{F} because of linearity in the first slot;
- $\langle 0, v \rangle$ and $\langle v, 0 \rangle = 0$ because linear maps take 0 to 0.
- The *norm* $\|v\| := \sqrt{\langle v, v \rangle}$. It has properties

- $\|v\| = 0$ iff $v = 0$, and
- $\|\lambda v\| = |\lambda|\|v\|$.
- Two vectors are *orthogonal* if their inner product is 0. Therefore, 0 is orthogonal to any vector and 0 is the only vector that is orthogonal to itself.
- Introducing these concepts provide us with a list of useful equations and inequalities about *inner product spaces* (vector spaces endowed with an inner product) that we know as facts for the Euclidean space.
 - The Pythagorean theorem now holds for orthogonal vectors $u, v \in V$:

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

- The method of orthogonal decomposition is extended to inner product spaces.
- The well-known Cauchy-Schwarz inequality states that $|\langle u, v \rangle| \leq \|u\|\|v\|$, which reaches equality iff $u = \lambda v$ for some scalar λ .
- The triangle inequality
- The parallelogram equality

6.2 Orthonormal Bases

6.3 Orthogonal Complement and Minimization Problems

7 Operators on Inner Product Spaces

7.1 Self-Adjoint and Normal Operators

7.2 The Spectral Theorem

7.3 Positive Operators and Isometries

7.4 Polar Decomposition and Singular Value Decomposition

8 Operators on Complex Vector Spaces

8.1 Generalized Eigenvectors and Nilpotent Operators

8.2 Decomposition of an Operator

8.3 Characteristic and Minimal Polynomials

8.4 Jordan Form

9 Operators on Real Vector Spaces

9.1 Complexification

9.2 Operators on Real Inner Product Spaces

10 Trace and Determinant

10.1 Trace

10.2 Determinant