Department of Computer Science and Engineering, NIT Calicut

Lecture 1: Basic Algebraic Structures

Prepared by: K Murali Krishnan

These notes assume that the reader has some familiarity with the notions of groups, rings fields and vector spaces. The definitions are stated here only for fixing the notation and exercises list out elementary facts which the reader is expected to know before proceeding further. Standard facts about matrices and determinants will be used without explanation.

Notation

Let $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$ denote the set of integers, rationals, reals and complex numbers respectively. Let $\mathbf{N} = \{0, 1, 2, ..\}$. We will use the notation $\mathbf{M_n(R)}$, $\mathbf{M_n(Q)}$, $\mathbf{M_n(C)}$ to denote the set of $n \times n$ matrices with real, rational and complex entries.

Groups, Rings Fields and Vector Spaces

Definition 1. *A* **monoid** $(G, .)$ *is a (non empty) set G together with an associative binary operator "." on G having an identity element (denoted by 1 or sometimes e). If "." is commutative, G will be called a* **commutative monoid**. *G is a* **group** *if in addition every element in G has an inverse. A commutative group is called an* **Abelian group**.

Exercise 1. *Find the category to which $(\mathbf{Z}, +)$, $(\mathbf{Z}, .)$, $(\mathbf{Q}, +)$, $(\mathbf{Q}, .)$, $(\mathbf{N}, +)$, $(\mathbf{N}, .)$ belong to where, "+" and "." represent standard addition and multiplication. What about $(Q \setminus \{0\}, .)$? and $(N \setminus \{0\}, .)$?*

Example 1. *$(M_n(X), +)$ for $X \in \{\mathbf{Q}, \mathbf{R}$ or $\mathbf{C}$ $\}$ and "+" the standard matrix addition is an Abelian group with zero matrix 0 as identity. $(M_n(X), .)$ for $X \in \{\mathbf{Q}, \mathbf{R}$ and $\mathbf{C}\}$ and "." the standard matrix multiplication is a (non-commutative) monoid with the $n \times n$ identity matrix I_n as identity. However the set $GL_n(X)$ consisting of non-singular $n \times n$ matrices over X forms a (non-Abelian) group with respect to multiplication.*

Definition 2. *A set $(R, +, .)$ with two operators is a* **ring (with unity)** *if $(R, +)$ is an Abelian group, $(R, .)$ is a monoid and "." distributes over "+". A ring R is a* **commutative** *if $(R, .)$ is a commutative monoid. A commutative ring R is a* **field** *if $(R \setminus \{0\}, .)$ is an Abelian group. Normally 0 and 1 are used to represent the additive and multiplicative identities.*

Exercise 2. *Which among $(\mathbf{Z}, +, .)$, $(\mathbf{N}, +, .)$, $(\mathbf{Q}, +, .)$ are rings?. Which among them are fields?*

Example 2. *$(M_n(\mathcal{R}), +, .)$ is a non-commutative ring with unity (identity matrix I_n).*

Vector Spaces

Definition 3. *An Abelian group $(V, +)$ is a vector space over a field F if there is* scalar multiplication *function "." from $F \times V$ to V satisfying $(a+b)v = av + bv$, $a(bv) = (ab)v$, $1v = v$, $a(v+w) = av + aw$ for all $a, b \in F$ and $v, w \in V$. Normally we write $V(F)$ to denote a vector space V over field F.*

Example 3. $\mathbf{R^n}$ *over* $\mathbf{R}$ *or* $\mathbf{Q}$ *(but not* $\mathbf{C}$ *– why?) is a vector space with addition and scalar multiplication defined in the standard way. So is* $\mathbf{C^n}$ *over* $\mathbf{R}$, $\mathbf{Q}$ *or* $\mathbf{C}$.

Example 4. *If F is any field, the set F^n consisting of n tuples over F is a vector space over F where multiplication of a vector with a scalar is defined (in the standard way) as component-wise multiplication. $M_n(\mathbf{X})$ is a vector space over X for $X \in \{\mathbf{Q}, \mathbf{R}, \mathbf{C}\}$. In general, if T is any set and F any field, then the set of functions from T to X (denoted by X^T) is a vector space over F with scalar multiplication defined in the standard way as $(\alpha f)(x) = \alpha f(x)$. The previous examples are special cases of this general case (how?).*

Example 5. *If F is a field, the set $F[x]$ of polynomials with coefficients in F is a vector space over F.*

Subspaces

Definition 4. *A subset V' of a vector space $V(F)$ is called a* **subspace** *if $V'(F)$ is a vector space.*

Example 6. *Consider $F[x]$ consisting of polynomials with coefficients in F. Consider $xF[x]$ which are polynomials with no constant term. It is easy to see that $xF[x]$ is a subspace of $F[x]$ over F. In general x may be replaced in this example with any $g(x) \in F[x]$.*

Example 7. *Consider* $\mathbf{R^2}$ *the two dimensional Cartesian place. Any line through the origin $\{(x, y) \in \mathbf{R^2} : (ax+by = 0)\}$ for any $a, b \in \mathbf{R}$ is a subspace. This subspace consists of the line through the origin perpendicular to the vector (a, b). The whole R^2 and the single point $(0, 0)$ are trivial subspaces. In general, in R^n, the (hyper) plane through the origin perpendicular to the vector $(a_1, a_2, ...a_n)$ will be the subspace defined by $a_1x_1 + a_2x_2 + ... + a_nx_n = 0$.*

Example 8. *The set of all $n \times n$ real matrices with determinant ± 1 denoted by $SL_n(\mathbf{R})$ (called* orthogonal matrices*) is a subgroup of $GL_n(\mathbf{R})$ with respect to multiplication.*

Exercise 3. *Suppose $V(F)$ is a vector space, show that $V' \subseteq V$ is a subspace if and only if for each $v, w \in V'$, $av + bw \in V'$ for any $a, b \in F$.*

Exercise 4. *Let $S = \{v_1, v_2, ..v_m\}$ be vectors in a vector space $V(F)$. Define $span(S) = \{a_1v_1 + a_2v_2 + ... + a_mv_m : a_1, a_2, ..a_m \in F\}$. Show that $span(S)$ is a subspace of V. Show that a span of a non-zero vector (x, y, z) in $\mathbf{R}^3(\mathbf{R})$ is a line through the origin. Show that two points (x, y, z) and (x', y', z') spans a plane if and only if $(0, 0, 0), (x, y, z)$ and (x', y', z') are not on the same line.*

Definition 5. *A Set of vectors S is* linearly dependent *if there are distinct vectors $v_1, v_2..., v_n$ in S and scalars $a_1, a_2, ..., a_n$ in F, not all zero satisfying $a_1v_1 + a_2v_2 + ... + a_nv_n = 0$. We follow the convention that $\emptyset$ is linearly independent and $\{0\}$ linearly dependent.*

A set of vectors S is linearly dependent if S is not linearly independent. That is, whenever $a_1v_1 + a_2v_2 + ... + a_nv_=0$ for distinct $v_1, v_2..., v_n \in S$ then $a_1 = a_2 = ... = a_n = 0$.

Example 9. *The vectors* $v_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *and* $v_2 \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ *are linearly dependent in* $\mathcal{R}^2$ *as* $2v_1 - v_2 = 0$. *The vectors* $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and* $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *are linearly independent.*

Example 10. *In general, the vectors* $e_1 = [1, 0, ..., 0]^T$, $e_2 = [0, 1, 0, ..., 0]^T$ $e_n[0, 0, ..., 1]^T$ *are linearly independent in* $\mathcal{R}^n$. *Moreover,* $span(\{e_1, e_2, ..., e_n\}) = \mathcal{R}^n$.

Example 11. $\{1, x, x^2, ..., x^n...\}$ *forms a linearly independent set in vector space* $F[x]$ *for any field* F. *The span of the set is the whole* $F[x]$.

Let $S = \{v_1, v_2, ..v_m\}$ be vectors in a vector space $V(F)$. Recall that $span(S) = \{a_1v_1 + a_2v_2 + ... + a_mv_m : a_1, a_2, ..a_m \in F\}$ is a subspace of V. $Span(S)$ is essentially the set of vectors expressible as finite linear combinations of vectors in S. The following lemma says that a set of vectors in linearly dependent if and only if one of the vectors is the span of the remaining.

Lemma 1. *A set of vectors* $v_1, v_2, ..., v_n$ *in a vector space* $V(F)$ *is linearly dependent if and only if for some* $k \leq n$, $v_k \in span(v_1, v_2, ..v_{k-1})$.

Proof. Let k be the smallest index such that $v_1, v_2, ..v_k$ are linearly dependent (why should such k exist?). Then, there exist $a_1, a_2, ..., a_k$ such that $a_1v_1+a_2v_2+...+a_kv_k = 0$. Moreover, $a_k \neq 0$ (why?). Hence $v_k = -(a_1/a_k)v_1 - (a_2/a_k)v_2 + ... - (a_{k-1}/a_k)v_{k-1}$. Converse is easy (why?). $\square$

In this lecture we will develop some elementary theory about vector spaces. Let $V(F)$ be a vector space over field F.

Definition 6. *A set S of vectors in $V(F)$ forms a* **basis** *for V if S is linearly independent and $span(S) = V$.*

Example 12. *It is easy to see that $v_1 = [x, y]^T$ and $v_2 = [x', y']^T$ forms a basis of $\mathcal{R}^2$ whenever they do not fall on a line passing through the origin.*

Lemma 2. *If $\{x_1, x_2, ..., x_n\}$ spans V and $\{y_1, y_2, .., y_m\}$ is a linearly independent set, the $m \leq n$. That is, the size of the largest independent set cannot exceed the size of the smallest spanning set for V (whenever there exists a finite set of vectors that span V).*

Proof. Since $y_m \in span\{x_1, x_2, ..., x_n\}$, the set $\{y_m, x_1, x_2, ..., x_n\}$ is linearly dependent. By previous lemma, there must be some x_i such that $x_i \in span\{y_m, x_1, x_2, .., x_{i-1}\}$. Hence we can eliminate x_i from the set and $\{y_m, x_1, x_2, .., x_{i-1}, x_{i+1}, ...x_n\}$ will be a spanning set. Now we add y_{m-1} to this set and remove another $x_{i'}$ from the resultant set and still get a spanning set. If we continue this process, $x_i s$ cannot be finished before all $y_j s$ are added for otherwise we will have $y_k, y_{k+1}, ..., y_m$ will be a spanning set for some $k > 1$ and this will be contradiction as then y_1 will be in the span of $y_k, y_{k+1}, ..., y_m$. Hence $n \geq m$. □

We are ready to prove the main theorem:

Theorem 1. *If V has a finite basis, then any two basis of V the same number of elements. This number is called the* **dimension** *of V.*

Proof. Let S and T be two (finite) basis for V. Since S is spanning and T linearly independent, we have $|S| \geq |T|$ by lemma above. Since T is spanning and S linearly independent, $|T| \geq |S|$. Hence $|S| = |T|$. □

V is said to be a finite dimensional if it has a finite basis. The dimension of V is denoted by $Dim(V)$.

Theorem 2. *Let $\{v_1, v_2, ..., v_n\}$ be a basis for a FDVS $V(F)$. Then for each $v \in V$, there exists unique $a_1, a_2, ..., a_n \in F$ such that $v = a_1v_1 + a_2v_2 + ... + a_nv_n$. $a_1, a_2, ..., a_n$ are called the* **coordinates of** v **with respect to basis** $v_1, v_2, ..., v_n$.

Proof. Clearly $a_1, a_2, ..., a_n$ must exist as $\{v_1, v_2, ..., v_n\}$ spans V. Suppose $v = a_1v_1 + a_2v_2 + ... + a_nv_n = b_1v_1 + b_2v_2 + ... + b_nv_n$, then $(a_1 - b_1)v_1 + (a_2 - b_2)v_2 + ... + (a_n - b_n)v_n = 0$. It follows from linear independence of $\{v_1, v_2, ..., v_n\}$ that $a_i = b_i$ for each i. □

To construct a basis for a FDVS $V(F)$, we can start with any vector v_1, pick v_2 outside $span(v_1)$, pick v_3 outside $span(v_1, v_2)$ and so forth. The process must terminate in finite number of steps as otherwise $v_1, v_2, ...v_n$ will be an infinite linearly independent set contradicting the finite dimensionality of V. (why?). Similarly, if W is a subspace of V, we can extend a basis of W to a basis of V exactly as above. (how?). Essentially we have proved the following:

Theorem 3. *Every finite dimensional vector space $V(F)$ has a basis. Moreover, basis for a subspace may be extended to a basis for V.*

The facts that every vector space has a basis and that any two basis have the same cardinality hold for arbitrary vector spaces - finite or infinite dimensional. However a study of infinite dimensional spaces is beyond our present scope of discussion.

Definition 7. *A map T from a vector space $V(F)$ to another $V'(F)$ (the field must be the same) is a homomorphism (or a* **linear transformation***) if $T(v + v') = T(v) + T(v')$ and $T(av) = aT(v)$ for all $v, v' \in V$ and $a \in F$. A bijective homomorphism is called an* **isomorphism**.

A isomorphism between two structures indicate that the two are identical except for a re-naming of elements (via the map).

Definition 8. *Let T be a linear transformation between two vector spaces V and V'. The image of the map T in V' is sometimes denoted by $img(T)$. The* **kernel** *of the map denoted by $ker(T)$ is the collection of elements in V that gets mapped to zero in V'. Dimension of $img(T)$ is called* **Rank(T)**. *Dimension of $ker(T)$ is called* **Nullity(T)**.

Example 13. *The map from $\mathcal{R}^3$ to $\mathcal{R}$ defined by $f(x, y, z) = x+y+z$ is a linear transformation. The map from $\mathcal{R}^2$ to itself which rotates each vector by θ degrees is a homomorphism. The action of the map on the point $\begin{bmatrix} x \\ y \end{bmatrix}$ is left multiplication by the matrix $\begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$*

Exercise 5. *Find the kernel and image of the maps above.*

Exercise 6. *Let F be any field and let $\alpha \in F$. The map Φ from $F[x]$ to F defined by $\Phi(f) = f(\alpha)$ is a homomorphism and is called the "evaluation map at α". prove that map is a vector space homomorphism. If $F = \mathcal{R}$ and $\alpha = 1$, what is the kernel and image? What if $\alpha = 0$? What if $\alpha = \pi$? (Hint: For the last part, you need to know the fact that there is no polynomial with real (in fact even complex) coefficients which has π as a root).*

Exercise 7. *Show that the kernel and image of linear transformations must be subspaces) of the respective spaces.*

Exercise 8. *Show that a linear is injective if and only if $ker(f) = \{0\}$ (or identity element for group homomorphisms). This is important as proving injectivity at zero suffices to prove injectivity of the map.*

Exercise 9. *Let T be a bijective linear transformation (isomorphism) between vector spaces $V(F)$ and $W(F)$. Let $b_1, b_2, ..., b_n$ be a basis of V. Show that $T(b_1), T(b_2), ..., T(b_n)$ is a basis of W. In particular, $dim(V) = dim(W)$.*

Theorem 4 (Rank Nullity Theorem)**.** *Let T be a linear transformation from $V(F)$ to $W(F)$. Then $Rank(T) + Nullity(T) = dim(V)$*

Proof. Let $b_1, b_2, ..., b_n$ be a basis of V. Let $Rank(T) = r$ and $Nullity(T) = k$. Clearly $0 \leq r, k \leq n$ Required to prove that $r + k = n$.

Consider $T(b_1), T(b_2), .., T(b_n)$. Clearly $Img(T) = Span\{T(b_1), ..., T(b_n)\}$ (why?). Since $Rank(T) = r$, exactly r, any maximal independent subset of these vectors must contain exactly r elements. Without loss of generality, assume $T(b_1), T(b_2)...T(b_r)$ are linearly independent. Let $c_1, c_2, .., c_k$ be a basis for $ker(T)$. We will show that the set $S = \{b_1, b_2, ..b_r, c_1, c_2, ..., c_k\}$ is a basis for V. Note that this proves the theorem (why?).

Suppose $\alpha_1 b_1 + ... + \alpha_r b_r + \beta_1 c_1 + ... + \beta_k c_k = 0$. Applying T and noting that $T(c_i) = 0$ for all $1 \leq i \leq k$, we have $T(\alpha_1 b_1 + ... + \alpha_r b_r) = \alpha_1 T(b_1) + ... + \alpha_r T(b_r) = 0$. Using linear independence of b_j for $1 \leq j \leq r$, we get $\alpha_1 = \alpha_2 = ... = \alpha_r = 0$. Using this fact in the first equation, we get $\beta_1 c_1 + ... + \beta_k c_k = 0$. Linear independence of c_i for $1 \leq i \leq k$ implies $\beta_i = 0$ for all $1 \leq i \leq k$. This establishes linear independence of S.

Now to show that S spans V, consider any vector $v \in V$. Since $T(v) \in Img(T)$, there must exist $\alpha_1, ... + ., \alpha_r$ in F such that $T(v) = \alpha_1 T(b_1) + ... + \alpha_r T(b_r) = T(\alpha_1 b_1 + ... + \alpha_r b_r)$. Hence $T(v - \alpha_1 b_1 - ... - \alpha_r b_r) = 0$ or $v - \alpha_1 b_1 - ... - \alpha_r b_r \in ker(T)$. Hence there must be $\beta_1, ..., \beta_k$ in F such that $v - \alpha_1 b_1 - ... - \alpha_r b_r = \beta_1 c_1 + ... + \beta_k c_k$. But this guarentees that v is in the span of S thereby completing the proof (why?).

$\square$

Exercise 10. *Let α be a real number. Consider the map Φ_α defined from $\mathcal{R}[x]$ to $\mathcal{R}$ defined by $\Phi_\alpha(f) = f(\alpha)$. For various values of α, what can you say about $\ker(\Phi_\alpha)$ and $img(\Phi_\alpha)$? What can you say about $Rank(\Phi_\alpha)$ and $Nullity(\Phi_\alpha)$ for various values of α?*

Exercise 11. *Let $b_1, b_2, .., b_n$ be a basis for $V(F)$. Suppose T is a linear map from V to $W(F)$ of dimension m. Show that for each choice of (not necessarily distinct vectors) $w_1, w_2, ..., w_n$ in W and setting $T(b_1) = w_1, T(b_2) = w_2, ..., T(b_n) = w_n$ we get a distinct linear transformation from V to W. Show that each linear transformation from V to W corresponds to a unique assignment of values for $T(b_1), t(b_2), ..., T(b_n)$ in W. This result is often stated as "fixing the image of the basis fixes the linear map".*

Exercise 12. *Let $V(F)$ be a vector space of dimension. Let $e_1 = [1, 0, .., 0]^T$, $e_2 = [0, 1, ..., 0]^T$, $e_n = [0, 0, ..., 1]^T$ be the standard basis of the vector space F^n. Let $b_1, b_2, .., b_n$ be any basis for $V(F)$. Define the map $T(b_1) = e_1$, $T(b_2) = e_2$,...,$T(b_n) = e_n$. Show that T is an isomorphism. It follows that every vector space of dimension n over F is isomorphic to F^n.*

Matrices

Let $V(F)$ have basis $b_1, b_2, ..., b_n$ and $W(F)$ have basis $c_1, c_2, ..., c_m$. Let T be a linear transformation from V to W. Let $T(b_1) = a_{11}c_1 + a_{12}c_2 + ... + a_{1m}c_m$. In dot product notation we write $T(b_1) = [c_1, c_2, ..., c_m][a_{11}, a_{12}, ..., a_{1m}]^T$. Similarly, let $T(b_2) = [c_1, c_2, ..., c_m][a_{21}, a_{22}, ..., a_{2m}]^T$,....., $T(b_n) = [c_1, c_2, ..., c_m][a_{n1}, a_{n2}, ..., a_{nm}]^T$.

Let $v = x_1, b_1 + x_2b_2 + ... + x_nb_n$. for some scalars $x_1, x_2, .., x_n$. By linearity of T,, $T(v) = x_1T(b_1) + x_2T(b_2) + ... + x_nT(b_n) = [T(b_1), T(b_2), ..., T(b_n)][x_1, x_2, ..., x_n]^T$ in dot product notation.

Noting that in dot product notation $T(b_i) = [c_1, c_2, ..., c_m][a_{i1}, a_{i2}, ..., a_{im}]^T$, we have in matrix notation:

$$\begin{bmatrix} T(b_1) & ... & T(b_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{bmatrix} = \begin{bmatrix} c_1 & ... & c_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & ... & a_{n1} \\ a_{12} & a_{22} & ... & a_{n2} \\ .. & .. & ... & .. \\ a_{1m} & a_{2m} & ... & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{bmatrix}.$$

Suppose $[y_1, y_2, ..., y_m]$ are the coordinates of $T(v)$ with respect to basis $c_1, c_2, ..., c_m$, then we have the relation:

$$\begin{bmatrix} y_1 \\ y_2 \\ .. \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & ... & a_{n1} \\ a_{12} & a_{22} & ... & a_{n2} \\ .. & .. & ... & .. \\ a_{1m} & a_{2m} & ... & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{bmatrix}.$$

Thus the matrix $A = \begin{bmatrix} a_{11} & a_{21} & ... & a_{n1} \\ a_{12} & a_{22} & ... & a_{n2} \\ .. & .. & ... & .. \\ a_{1m} & a_{2m} & ... & a_{nm} \end{bmatrix}$ is called the matrix of the linear transformation with respect to basis $b_1, b_2, ..., b_n$ and $c_1, c_2, .., c_m$. Conversely, it is easy to see that any $m \times n$ matrix will define a linear transformation for the basis of particular choice. Thus we see a correspondence between $m \times n$ matrices over the field F and linear transformations from V to W.

We have already seen that any n dimensional vector space over F is isomorphic to F^n. Hence, *once we fix a basis for V and W*, vectors from V correspond to elements in F^n, vectors in W correspond to elements in F^m and linear transformation from V to W correspond to $m \times n$ matrices over F. This correspondence draws matrices into the study of linear transformations.

In these lectures, we will be specific to the following special class of linear transformations.

Definition 9. *A* **(linear) operator** *on a vector space $V(F)$ is a linear transformation from V to itself.*

Once a(ny) basis for an n dimensional vector space V is fixed, each linear operator on V corresponds to a $n \times n$ square matrix. Thus, the set of operators on an n dimensional space V corresponds precisely to $M_n(F)$.

Exercise 13. *Let $b_1, b_2, ..., b_n$ be a basis for $V(F)$. Show that an operator T on V is bijective if and only if T is injective if and only if $T(b_1), T(b_2), ..., T(b_n)$ are linearly independent. Note that a linear transformation T is* **invertible** *if and only if T is bijective. Show that T^{-1} is also a linear operator from V to V. (why?).*

Let $b_1, b_2, ..., b_n$ be a basis of $V(F)$. We have already seen that the map $f : V \longrightarrow F^n$ defined by $f(b_1) = e_1, ..., f(b_n) = e_n$ is an isomorphism. With this identification, a vector $v = x_1b_1 + x_2b_2 + ... + x_nb_n$ may be identified with $[x_1, x_2, ..., x_n]^T \in F^n$ Now, let T be an operator in V. Then the matrix A of the map has coordinate vectors corresponding to $T(e_1), T(e_2), ...T(e_n)$ as columns (with our identification of e_i with b_i). In view of the above exercise, we see that T is invertible if and only if the columns of A are linearly independent. This in turn happens if and only if the space spanned by the columns of A is the whole of V (why?). This observation motivates the following definition:

Definition 10. *Let $A \in F^{n\times n}$ be an $n \times n$ matrix. $ColumnSpan(A)$ is defined as the subsace spanned by the columns of A. $RowSpan(A)$ is defined as the subspace spanned by the rows of A. The dimensions of the column and row space are called $RowRank(A)$ and $ColumnRank(A)$ of A.*

It follows from the previous discussion that an $n \times n$ matrix A over a field F is invertible if and only if $ColumnSpan(A) = F^n$. Since we A is invertible if and only if $det(A) \neq 0$, we have a correspondance between bijective linear operators and matrices in $GL_n(F)$.

Corollary 1. *$T : V \longrightarrow V$ is bijective (invertible) if and only if the matrix of T (with respect to any basis $b_1, b_2, .., b_n$) is non-singular.*

Basis Transformations

We study the effect of basis change on the coordinates of a vector. The matrix of an operator also changes when basis changes.

Let $B = b_1, b_2, ..., b_n$ and $C = c_1, c_2, ..., c_n$ be two basis for $V(F)$. Suppose we know the coordinates of vectors in S' wrt. those in S. i.e., let $c_1 = \alpha_{11}b_1 + \alpha_{12}b_2 + ... + \alpha_{1n}b_n$, $c_2 = \alpha_{21}b_1 + \alpha_{22}b_2 + ... + \alpha_{2n}b_n$,..., $c_n = \alpha_{n1}b_1 + \alpha_{n2}b_2 + ... + \alpha_{nn}b_n$. In matrix notation, $[c_1, c_2, ..., c_n] = [b_1, b_2, ..., b_n]Q$ where, $Q = \begin{bmatrix} \alpha_{11} & \alpha_{21} & ... & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & ... & \alpha_{n2} \\ .. & .. & ... & .. \\ \alpha_{1n} & \alpha_{2n} & ... & \alpha_{nn} \end{bmatrix}$

Since basis transformation is an isomorphism, Q must be invertible (why?). Thus we have $[b_1, b_2, .., b_n] = Q^{-1}[c_1, c_2, ..., c_n]$. Suppose now $v = x_1b_1 + x_2b_2 + ... + x_nb_n$ be a vector with coordinates $[x_1, x_2, ..., x_n]^T$ with respect to basis B. What will be the coordinates of v with respect to basis C? That is, we want to find out $[y_1, y_2, ..., y_n] \in F^n$ such that $v = [c_1, c_2, ..., c_n][y_1, y_2, ..., y_n]^T$. But $v = [b_1, b_2, ..., b_n][x_1, x_2, ..., x_n]^T = [c_1, c_2, ..., c_n]Q^{-1}[x_1, x_2, ..., x_n]^T$. Hence we have $[y_1, y_2, ..., y_n]^T = Q^{-1}[x_1, x_2, ..., x_n]^T$ giving the required relation between coordinate vectors. Q is called the matrix of basis change from B to C.

Example 14. *In $\mathcal{R}^2$, let v have coordinates $[1,1]^T$ w.r.t. the standard basis. To find its coordinates w.r.t. basis $c_1 = [1,1]^T$ and $c_2 = [1,0]^T$, we can see that $[c_1, c_2] = [e_1, e_2]Q$ where $Q = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Thus the new coordinates will be $Q^{-1}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.*

Now we take up the effect of basis change on the matrix of a linear operator on a FDVS. Let $B = \{b_1, b_2, ..., b_n\}$ and $C = \{c_1, c_2, ..., c_n\}$ be two basis for an FDVS $V(F)$. Let $[c_1, c_2, ..., c_n] = [b_1, b_2, ..., b_n]Q$. Let A be the matrix of a linear operator with respect to basis B. Let v be a vector in V whose coordinate vector w.r.t. basis B is $x = [x_1, x_2, .., x_n]^T$. It follows that the coordinates of v w.r.t. basis C will be $Q^{-1}x$.

Since A is the matrix of T w.r.t. basis B, coordinate vector of $T(v)$ w.r.t. basis B will be Ax. Hence the coordinate vector for $T(v)$ w.r.t. basis C will be $Q^{-1}Ax$.

Let A' be the matrix of T w.r.t. basis C. As v has coordinates $Q^{-1}x$ w.r.t. C and $T(v)$ has coordinates $Q^{-1}Ax$ w.r.t. C, action of A' on $Q^{-1}x$ must give $Q^{-1}Ax$. That is, we must have $A'Q^{-1}x = Q^{-1}Ax$. Hence we have $A'x = Q^{-1}AQx$. Since this must hold for all $x \in F^n$ as v was chosen arbitrary, we have $A' = Q^{-1}AQ$ as the matrix of T for the basis C.

Example 15. *T be the linear operator in $\mathcal{R}^2$ such that $T(\begin{bmatrix} 1 \\ 0 \end{bmatrix}) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $T(\begin{bmatrix} 0 \\ 1 \end{bmatrix}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ The matrix of T w.r.t. the standard basis is $\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$. If we change the basis to $\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$ then the matrix of basis change $Q = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Hence the matrix of T w.r.t this basis will be $Q^{-1}AQ$= $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$*

Exercise 14. *Consider the operator T in $\mathcal{R}^3$ given by $T(e_1) = e_1$, $T(e_2) = e_1 + e_2$, $T(e_3) = e_1 + e_2 + e_3$. What is the matrix of this map w.r.t. the basis $b_1 = e_1 + e_2$, $b_2 = e_2 + e_3$ and $b_3 = e_1 + e_3$. (Hint, work with the relationship between the basis vectors directly instead of going for matrix manipulation and note that coordinate vectors of $T(b_1)$, $T(b_2)$ and $T(b_3)$ in the basis $\{b_1, b_2, b_3\}$ forms the columns of the matrix to be computed).*

Exercise 15. *Consider the set $F_n[x]$ consisting of all polynomials of degree less than n over a field F. Let $\alpha_1, \alpha_2, ..., \alpha_n$ be elements in F. Consider the map $T(p(x)) = p(\alpha_1) + p(\alpha_2)x + ..., +p(\alpha_n)x^n$ in $F_n[x]$. What is the matrix of the map with respect to the basis $\{1, x, x^2, ..., x^n\}$? This matrix is called a Vandermone's matrix. Find the expression for the determinant of the matrix and show that the map is invertible if and only if $\alpha_1, \alpha_2, ..., \alpha_n$ are distinct elements in F. This means that interpolation of a degree $n-1$ polynomial is possible only if evaluation at n distinct points are given. Moreover interpolation problem reduces to matrix inversion.*

In the following, assume that V is a vector space of dimension n over a field F.

Linear Forms

Definition 11. *A linear map l from V to F is called a linear form. The set of all linear forms from V to F will be denoted by V^*.*

Untill stated otherwise, fix basis $(b_1, b_2, \cdots, b_n)$ of V. Since F is a one dimensional vector space over F with basis (1), the matrix of a linear form l w.r.t basis $[b_1, b_2, \cdots, b_n]$ of V and [1] of F will be a $1 \times n$ row vector $[l(b_1), l(b_2), \cdots, l(b_n)]$. Thus, each linear from defines a unique row vector of n scalars over F. Suppose $v = x_1b_1 + x_2b_2 + \cdots + x_nb_n$, then $l(v) = \begin{bmatrix} l(b_1) & ... & l(b_n) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{bmatrix}$ Thus, fixing a basis, we reduce evaluation of a linear from to a dot product computation. Consequenly, we will think of each linear form as row vector, once a basis is fixed. The application of a linear from on a vector (column vector) amounts to finding the dot product of the two vectors.

Exercise 16. *Let $(b_1, b_2, \cdots, b_n)$ be a basis of V. Let $l_1, l_2 \in \mathbf{L}(V)$. Show that for any $\alpha_1, \alpha_2 \in F$, $\alpha_1 l_1 + \alpha_2 l_2 \in V^*$. What will be matrix (row vector) of $\alpha_1 l_1 + \alpha_2 l_2$ w.r.t the basis $[b_1, b_2, \cdots, b_n]$ of V and [1] of F?*

Exercise 17. *Using Exercise 11, conclude that V^* is isomorphic to F^n.*

Since it is easy to see that $\alpha_1 l_1 + \alpha_2 l_2 \in V^*$ whenever $l_1, l_2 \in \mathbf{L}(V)$, $\mathbf{L}(V)$ is a vector space. To find a basis for V^*, consider the set of linear forms: $(l_1, l_2, \cdots, l_n)$ defined as: $l_1(b_1) = 1, l_1(b_i) = 0$ for $i \neq 1$, $l_2(b_2) = 1$, $l_2(b_i) = 0$ for $i \neq 2$, $\cdots$, $l_n(b_n) = 1$, $l_n(b_i) = 0$ for $i \neq n$. Given a vector $v = x_1b_1 + x_2b_2 + \cdots + x_nb_n$, it is easy to see that $l_i(v) = x_i$ (why?). Thus, the application of the function l_i simply extracts the i^{th} component of v, with respect to the basis $(b_1, b_2, \cdots, b_n)$.

Exercise 18. *Show that the matrix of l_i with respect to the basis $[b_1, b_2, \cdots b_n]$ of V and [1] of F is e_i^T where e_i is the i^{th} standard basis vector.*

Theorem 5. *$l_1, l_2, \cdots l_n$ is a basis of V^*.*

Proof. Suppose $\alpha_1 l_1 + \alpha_2 l_2 + \ldots + \alpha_n l_n = 0$ for some $\alpha_1, \alpha_2, \cdots, \alpha_n \in F$. (Note that this a an expression involving linear combinations of functions and hence denotes a function. A function is zero if and only if it yields zero on any argument). Evaluation of the left side at the vector b_1 yields $\alpha_1 l_1(b_1) + \alpha_2 l_2(b_1) + \ldots + \alpha_n l_n(b_1) = 0$. Since $l_i(b_1) = 0$ for $i \neq 1$, this reduces to $\alpha_1 l_1(b_1) = 0$. Since $l_1(b_1) = 1$ by definition, the only possibility is that $\alpha_1 = 0$.

Similarly $\alpha_i = 0$ for all $1 \leq i \leq n$. This shows linear independence of $l_1, l_2, \cdots l_n$. Suppose $l \in \mathbf{L}(V)$. Set $\beta_1 = l(b_1), \beta_2 = l(b_2), \cdots, \beta_n = l(b_n)$. Note that $\beta_1, \beta_2, \cdots, \beta_n \in F$. We claim that $l = \beta_1 l_1 + \beta_2 l_2 + \ldots + \beta_n l_n$. To prove this, it suffices to prove that on each basis vector b_i, $l(b_i)$ evaluates to the same value as $(\beta_1 l_1 + \beta_2 l_2 + \ldots + \beta_n l_n)(b_i)$. Now $l(b_1) = \beta_1$, whereas $(\beta_1 l_1 + \beta_2 l_2 + \ldots + \beta_n l_n)(b_1) = \beta_1$ as $l_i(b_1) = 0$ when $i \neq 1$ and $l_1(b_1) = 1$. Similarly, the claim holds for each b_i. $\square$

Definition 12. *V^* is called the dual space of V. Given basis $b_1, b_2, \cdots, b_n$, the corresponding basis $l_1, l_2 \ldots l_n$ of V^* as defined above is called the dual basis of V^* corresponding to the basis $b_1, b_2 \ldots b_n$ in V.*

Exercise 19. *Let V, W be vector spaces over field F. Let $\mathbf{L}(V, W)$ be the set of all linear transformations from V to W. Let $b_1, b_2, \cdots, b_n$ be a basis of V and $c_1, c_2, \cdots, c_m$ be a basis of W.*

1. *Show that $\mathbf{L}(V, W)$ is a vector space.*

2. *Define the family of linear tranformations $l_{i,j} : V \mapsto W$ as follows. $l_{i,j}(b_i) = c_j$. $l_{i,j}(b_k) = 0$ for $i \neq k$, $1 \leq i \leq n$, $1 \leq j \leq m$. What is the matrix of $l_{i,j}$ with respect to the basis $b_1, b_2, \cdots, b_n$ of V and $c_1, c_2, \cdots c_m$ of W?*

3. *Show that the set $\{l_{i,j} : 1 \leq i \leq n, 1 \leq j \leq m\}$ is a basis for $\mathbf{L}(V, W)$. What can you conclude about the dimension of $\mathbf{L}(V, W)$?*

Null Spaces

Let $L \subseteq V^*$. We define the **null space** of L denoted by $L^0 = \{v \in V : l(v) = 0 \text{ for all } l \in L\}$ The null space of L.

We start with some observations.

Exercise 20. *Let $L \subseteq V^*$ and let $S \subseteq V$. Let $span(L)$ be the subspace of V^* spanned by L. Define $S^0 = \{l \in V^* : l(v) = 0 \text{ for all } v \in S\}$ Show that:*

1. *$L^0 = span(L)^0$.*

2. *L^0 is a subspace of V.*

3. *$S^0 = span(S)^0$.*

4. *S^0 is a subspace of V^*.*

Exercise 21. *Find L^0 for:*

1. *$L = \{[1, 1, 1], [0, 0, 1]\}$ in $(\mathbf{R}^3)^*$ (assuming standard basis).*

2. *$L = \{[0, 1, 1, 0], [1, 0, 0, 1]\}$ in F^4 where F is the binary field F_2 containing only elements $\{0, 1\}$ where addition is the XOR operation and multiplication is the AND operation.*

Let L be a subspace of V^*. The next theorem connects the dimension of L with the dimension of L^0.

Theorem 6 (Duality Theorem)**.** *Let V be a finite dimensional vector space of dimension n over field F. Let L be a subspace of V^* Then $Dim(L^0) = n - Dim(L)$.*

Proof. By Exercise 20, L^0 is a subspace of V. Let $Dim(L^0) = k$. Let $b_1, b_2, \cdots, b_k$ be a basis for L^0. Extend this to a basis $b_{k+1}, \cdots b_n$ of V. Consider the dual basis $l_1, l_2 \ldots l_n$ of V^*. It suffices to prove that $l_{k+1}, l_{k+2} \ldots l_n$ spans L (why?). Let $l \in L$. Let $l = \alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n$ for some scalars $\alpha_1, \alpha_2 \ldots \alpha_n \in F$. (why should such scalars exist?). Since $l(b_1) = l(b_2) = \ldots l(b_k) = 0$ (why?), we have $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$ (why?), completing the proof. (why?) □

Corollary 2 (Rank Theorem)**.** *Let $A \in F^{n \times n}$. Then $RowRank(A) = ColumnRank(A)$.*

Proof. Consider the map defined by A with repsect to the standard basis on F^n. By Exercise 20, $ker(A) = RowSpan(A)^0$ (why?). Consequently, by Duality theorem we have $Nullity(A) = n - RowRank(A)$. By the Rank Nullity theorem, we have $Nullity(A) = n - ColumnRank(A)$. □

Lecture 5: Eigen Values and Vectors

Prepared by: K Murali Krishnan

For the rest of this lecture, let T be a linear operator defined on a vector space V of dimension n over a field F. A non-zero vector $v \in V$ is an Eigen vector of T if there exists $\lambda \in F$ such that $T(v) = \lambda v$. Note that the action of T on an Eigen vector v results in a vector collinar to v. A scalar $\lambda \in F$ is an Eigen value of T if there exists $v \in V$, $v \neq 0$ such that $T(v) = \lambda v$.

Lemma 3. *If $v_1, v_2, \ldots, v_k$ are Eigen vectors of T with distinct Eigen values $\lambda_1, \lambda_2 \ldots \lambda_k$, then $v_1, v_2 \ldots v_n$ are linearly independent.*

Proof. Let $\{v_1, v_2 \ldots v_r\}$, $r \leq k$ be chosen from $v_1, v_2 \ldots v_k$ such that no proper subset of $\{v_1, v_2 \ldots v_r\}$ is linearly dependent. Let $\alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_r v_r = 0$ Since $T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r) = 0$, we have: $\lambda_1 \alpha_1 + \lambda_2 \alpha_2 v_2 + \cdots \lambda_r \alpha_r v_r = 0$ Multiplying the first equation with λ_1 and subtracting from the second, we get: $(\lambda_2 - \lambda_1)\alpha_2 v_2 + \cdots + (\lambda_n - \lambda_1)\alpha_n v_n = 0$, Which contradicts the assumption that no proper subset of $\{v_1, v_2 \ldots v_r\}$ is linearly dependent. $\square$

An Eigen basis for T is a basis of V consisting of Eigen vectors of V. Let λ be an Eigen value of T. $E_\lambda = \{v \in V : T(v) = \lambda v\}$ is called the Eigen space associated with the Eigen value λ. Clearly E_λ is a subspace of V (why?).

Exercise 22. *If $b_1, b_2 \ldots b_n$ is a basis of Eigen vectors of T with Eigen values $\lambda_1, \lambda_2 \ldots \lambda_n$ respectively (not necessarily distinct), Show that the matrix of T with respect to the basis $b_1, b_2 \ldots b_n$ is a diagonal matrix with $\lambda_1, \lambda_2, \cdots, \lambda_n$ in the diagonal.*

The above exercise shows that a basis of Eigen vectors, if found, would form a convenient cordinate system to study a linear operator, because with respect to that basis, the matrix of the operator becomes a diagonal matrix.

Note that v is an Eigen vector of T with Eigen value λ if and only if $(T - \lambda I)v = 0$, where I is the identity operator on V defined by $I(v) = v$ for each $v \in V$. Equivalently, v is an Eigen vector of T if and only if v is a member of the nullspace of the operator $T - \lambda I$. Consequenly, λ is an Eigen value of T if and only if the null space of $T - \lambda I$ contains at least one non-zero vector.

In what follows, assume that $[b_1, b_2, \ldots, b_n]$, $[c_1, c_2, \ldots, c_n]$ be bases of V with $[c_1, c_2, \ldots c_n] = [b_1, b_2, \ldots, b_n]Q$ for some $n \times n$ matrix Q over F. Let A be a matrix of T with respect w.r.t $[b_1, b_2, \ldots, b_n]$, then we have already seen that the matrix of T w.r.t $[c_1, c_2, \ldots c_n]$ will be $A' = Q^{-1}AQ$. We have further seen that a vector v with cordinates $[x_1, x_2, \ldots, x_n]^T$ w.r.t $[b_1, b_2, \ldots, b_n]$ will have cordinates $Q^{-1}[x_1, x_2, \ldots, x_n]^T$ w.r.t $[c_1, c_2, \ldots, c_n]$. if v is an Eigen value of T with Eigen value λ, $A[x_1, x_2, \ldots, x_n]^T = \lambda[x_1, x_2, \ldots, x_n]^T$, or equivalently, $(A - \lambda I)[x_1, x_2, \ldots, x_n] = 0$.

Exercise 23. *Show that $[x_1, x_2, \ldots, x_n]$ is an Eigen vector of A with Eigen value λ if and only if $Q^{-1}[x_1, x_2, \ldots x_n]$ is an Eigen vector of A' with Eigen value λ.*

Exercise 24. *Show that:*

1. $det(A - \lambda I) = det(A' - \lambda I)$ *for any* $\lambda \in F$.

2. $det(A - \lambda I) = 0$ *if and only if* λ *is an Eigen value of* T.

Note that $det(A - \lambda I) = 0$ is a polynomial equation (in the unknown λ) and can be solved by standard techniques. The polynomial $det(A - \lambda I)$ is called the **characteristic polynomial** and is denoted by $\chi_T(\lambda)$ of T (we write $\chi(\lambda)$ when T is clear from the context). Observe that $det(A' - \lambda I) = det(Q^{-1}AQ - Q^{-1}(\lambda I)Q) = det(Q^{-1})det(A - \lambda I)det(Q) = det(A - \lambda I)$ and hence the characteristic polynomial $\chi(\lambda)$ of T does not change when we change the basis. The Eigen values are the roots of $\chi(\lambda)$.

Exercise 25. *Show that* $\chi(0) = det(A)$. *Hence the constant term of the characteristic polynomial gives the value of the determinant. In particular conclude that similar matrices have the same determinant.*

Exercise 26. *If* $\chi(\lambda) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ *for some* $\lambda_1, \lambda_2, \ldots, \lambda_n \in F$ *(not necessarily distinct), show that:*

1. $det(A) = (-1)^n \prod_{i=1}^n \lambda_i$. *Thus, whenever* $\chi(\lambda)$ *factorizes completely in* F *into linear factors the product of the roots (Eigen values) yields* $(-1)^n det(A)$.

2. *Show that the sum of the Eigen values is equal to the coefficient of* x^{n-1} *of* $\chi(T)$.

Exercise 27. *Show that* T *is not bijective if and only if* 0 *is not an Eigen value of* T.