MATH 325 Q1: LINEAR ALGEBRA III, PART 1

Throughout this lecture F denotes either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers. Vector spaces are always assumed to be of finite dimension, unless otherwise stated.

We denote by $\mathrm{M}_n(F)$ (respectively $\mathrm{M}_{m\times n}(F)$) the set of all matrices with coefficients in F of size $n \times n$ (respectively $m \times n$). We denote by $\mathrm{O}_{m\times n} \in \mathrm{M}_{m\times n}(F)$ the $m \times n$-matrix whose entries are all zero, and by $\mathrm{I}_n \in \mathrm{M}_n(F)$ the $n \times n$-identity matrix.

Complex Numbers

In mathematics people want all polynomial equations with real coefficients to have solutions. In particular, they want an existence of a solution of the equation $x^2+1=0$. To deal with the problem that this equation has no real solutions mathematicians invented "the imaginary number" i having the property $i^2 = -1$ or equivalently $i = \sqrt{-1}$. In this way they came in a natural way to the notion of complex numbers.

Definition. *A complex number is a formal sum $a+b\cdot i$ where $a, b \in \mathbb{R}$. The number a is called the real part and the number b is called the imaginary part. Two complex numbers $z_1 = a+b\cdot i$ and $z_2 = c+d\cdot i$ are equal if and only if $a=c$, $b=d$. The set of all such expressions is denoted by $\mathbb{C}$.*

Remark. *We may identify any real number a with a formal sum $a+0\cdot i$. In this way we may view the set $\mathbb{R}$ of all real numbers as a subset of $\mathbb{C}$.*

The following operations on $\mathbb{C}$ are clearly well-defined.

(1) **Addition:**

$$(a+b\,i)+(c+d\,i) = (a+b)+(c+d)\,i;$$

(2) **Subtraction:**

$$(a+b\,i)-(c+d\,i) = (a-c)+(b-d)\,i;$$

(3) **Multiplication:**

$$(a+b\,i)(c+d\,i) = (ac-bd)+(ad+bc)\,i;$$

Multiplication is commutative: $z_1\cdot z_2 = z_2\cdot z_1$; associative: $(z_1z_2)z_3 = z_1(z_2z_3)$. The real number $1 = 1+0\cdot i$ is the multiplicative identity.

(4) **Complex conjugation:** if $z = a+b\,i$ then the complex conjugation $\bar{z}$ of z is

$$\bar{z} = a-b\,i.$$

One easily checks that $z\cdot\bar{z} = a^2+b^2 \in \mathbb{R} \subset \mathbb{C}$.

(5) **The absolute value:** If $z = a+b\,i$ then the absolute value of z (or modulus of z) is the real number

$$|z| = \sqrt{a^2+b^2}.$$

(6) **Existence of multiplicative inverse:** if $z \in \mathbb{C}$ is nonzero then it has the multiplicative inverse; i.e. there exists $w \in \mathbb{C}$ such that $z \cdot w = 1$. Formula: if $z = a + b\,i$ then

$$w = \frac{\bar{z}}{|z|^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} \cdot i.$$

This complex number w is usually denoted by z^{-1}.

(7) **Division:** if $z, w \in \mathbb{C}$ and $w \neq 0$ then

$$\frac{z}{w} := z \cdot w^{-1}.$$

Geometric presentation of complex numbers

Given a complex number $z = a + b\,i$ one can associate the ordered pair (a, b) of real numbers which is geometrically represented by a point (or vector) in xy-plane with coordinates a and b respectively. The geometric picture then suggests that z can be presented in the form

$$z = |z|(\cos\,\phi + \sin\phi \cdot i)$$

which is called a polar form of z. The angle ϕ in this formula is called an argument of z. It is not unique. It is defined up to multiples of 2π. However, there is only one angle satisfying

$$-\pi < \phi \leq \pi$$

which is called the principal argument of z and is denoted by $\arg(z)$.

Example. If $z = 1 - \sqrt{3} \cdot i$ then

$$z = 2(\cos(-\frac{\pi}{3}) + \sin(-\frac{\pi}{3}) \cdot i).$$

Presenting complex numbers in geometric form helps us to multiply them, divide, take powers and extract roots: if z_1 and z_2 have arguments ϕ and ψ then

$$z_1 \cdot z_2 = |z_1||z_2|(\cos(\phi + \psi) + \sin(\phi + \psi)\,i).$$

$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|}(\cos(\phi - \psi) + \sin(\phi - \psi)\,i).$$

Also,

$$z_1^n = |z_1|^n(\cos(n\phi) + \sin(n\phi))$$

and

$$\sqrt[n]{z} = \sqrt[n]{|z|}(\cos(\phi/n) + \sin(\phi/n)\,i).$$

Remark. *By definition, $\sqrt[n]{z}$ is a complex number w such that $w^n = z$. In other words, it is a solution of the equation $x^n = z$. It is worth mentioning that such solution is not unique. Indeed, if ξ_n is a root of* *unity of degree n, i.e. $\xi^n = 1$ then $\widetilde{w} = w \cdot \xi_n$ also has the property $\widetilde{w}^n = z$. Conversely, if w_1 is another solution of the equation $x^n = z$ then $w_1 = w \cdot \xi_n$ for some root of unity of degree n.*

Remark. *For a given n we have exactly n roots of unity of degree n. All of them are given by the formula*

$$\cos\left(\frac{2\pi\, k}{n}\right) + \sin\left(\frac{2\pi\, k}{n}\right) i$$

where $k = 0, 1, \dots, n-1$.

Vector spaces

Recall in class the definitions of vector spaces over a field F and a vector subspace. Throughout F is either $\mathbb{R}$ or $\mathbb{C}$.

Main Example. Let

$$V = F^n = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \;\mid\; a_i \in F \right\}.$$

The addition is given by componentwise and multiplication by a scalar is given by multiplication of all components.

Definition. *Let V, U be vector spaces over F. A map $f : V \to U$ is called linear (linear transformation or linear operator) if it preserves linear combinations, i.e. for any two vectors $u, v \in V$ and scalars $a, b \in F$ one has $f(au + bv) = af(u) + bf(v)$. A linear map $f : V \to V$ is called a (linear) endomorphism of V.*

Definition. *A linear map $f : V \to U$ is called an isomorphism if f is bijective.*

Definition. *Let $f : V \to U$ be a linear map. The set*

$$\{v \in V \mid f(v) = 0 \}$$

is called the kernel of f and is denoted by $\operatorname{Ker} f$. The subset

$$\operatorname{Im} f = \{f(v) \mid v \in V\}$$

of U is called the image of f.

Remark. *It is almost obvious that $\operatorname{Ker} f = 0$ if and only if f is injective and $\operatorname{Im} f = U$ if and only if f is surjective.*

Definition. *Let V be a vector space over F. A subset $\{v_1, \dots, a_n\} \subset V$ is called linear independent if from $a_1 v_1 + \cdots + a_n v_n = 0$ where $a_1, \dots, a_n \in F$ it follows that $a_1 = \cdots = a_n = 0$. It is called a basis of*

V if it is in addition is a system of generators, i.e. every vector $v \in V$ can be written as a linear combination of $v_1, \dots, v_n$.

Definition. *Let V be a vector space over F and let $f : V \to V$ be an endomorphism. A nonzero vector $v \in V$ is called an eigenvector for f if there exists a scalar $\lambda \in F$ such that $f(v) = \lambda v$. This scalar λ is called an eigenvalue of f. Often one says that v is a λ-eigenvector.*

Remark. *It is easy to see that if v is a λ-eigenvector for f then for any scalar $b \in F$ the vector $\tilde{v} = bv$ is also a λ-eigenvector. Thus all nonzero vectors in 1-dimensional subspace $U = \langle v \rangle \subset V$ spanned by v are λ-eigenvectors for f. Conversely, if $U \subset V$ is a 1-dimensional subspace stable with respect to f then every nonzero vector $v \in$ is an eigenvector for f. Indeed, since U is f-stable then $f(v) \in U$, hence $f(v) = av$ for some scalar $a \in F$. We used here the fact that U is spanned by v.*

Linear maps and matrices

We briefly recall here the relation between linear maps and matrices.

Let $\alpha : V \longrightarrow W$ be a linear map between the (finite dimensional) vector spaces V and W. Let $v_1, \dots, v_n$ and $w_1, \dots, w_m$ be bases of V and W, respectively. Then there exists unique elements $a_{ij} \in F$, $1 \le i \le m$, $1 \le j \le n$, such that

$$\alpha(v_j) = \sum_{i=1}^{m} a_{ij} w_i \,. \tag{1}$$

The $m \times n$-matrix

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

is the matrix of α with respect to these bases.

If $V = W$ and $v_i = w_i$ we say also that A is the matrix of α with respect to the basis $v_1, \dots, v_n$.

Note that the matrix A depends on choosen bases of V and W. However bases are not unique. Our next goal is to see what happens with A if we choose other bases of V and W.

Let $v'_1, \dots, v'_n$ and $w'_1, \dots, w'_m$ are other bases of V and W, respectively. Then there are unique elements s_{ij} and t_{ij} of F, such that

$$v'_i = \sum_{h=1}^{n} s_{hi} v_h \qquad \text{and} \qquad w'_i = \sum_{h=1}^{m} t_{hi} w_h \,. \tag{2}$$

Similarly, there are $s'_{ij}, t'_{ij} \in F$, such that

$$v_i = \sum_{h=1}^{n} s'_{hi} v'_h \qquad \text{and} \qquad w_i = \sum_{h=1}^{m} t'_{hi} w'_h \,. \tag{3}$$

Inserting the equations (2) in (3) the linear independence of $v_1, \dots, v_n$ and of $w_1, \dots, w_m$ implies

$$\sum_{h=1}^{n} s_{lh} \cdot s'_{hi} \;=\; \sum_{h=1}^{m} t_{lh} \cdot t'_{hi} \;=\; \begin{cases} 1 & i = l \\ 0 & i \neq l \,, \end{cases}$$

or, in other words, the matrices $S = (s_{ij})$ and $T = (t_{ij})$ are invertible with inverses $S^{-1} = (s'_{ij})$ and $T^{-1} = (t'_{ij})$, respectively.

We use this now to compute the matrix $A' = (a'_{ij})$ of the linear map α with respect to the bases $v'_1, \dots, v'_n$ and $w'_1, \dots, w'_m$. We have

$$\begin{aligned} \alpha(v'_i) = \alpha\Big(\sum_{h=1}^{n} s_{hi} v_h\Big) \;&=\; \sum_{h=1}^{n} s_{hi} \alpha(v_h) && \text{by (2)} \\ &= \sum_{h=1}^{n} s_{hi} \sum_{l=1}^{m} a_{lh} w_l && \text{by (1)} \\ &= \sum_{h=1}^{n} s_{hi} \sum_{l=1}^{m} a_{lh} \sum_{j=1}^{m} t'_{jl} w'_j && \text{by (3)} \\ &= \sum_{j=1}^{m} \Big(\sum_{l=1}^{m} t'_{jl} \Big(\sum_{h=1}^{n} a_{lh} s_{hi} \Big) \Big) \cdot w'_j \,. \end{aligned}$$

Therefore $a'_{ij} = \sum_{l=1}^{m} t'_{jl} \Big(\sum_{h=1}^{n} a_{lh} s_{hi} \Big)$, and this is the ij-coefficient of the matrix $T^{-1} \cdot A \cdot S$. In other words, the matrix of α in bases $\{v'_1, \dots, v'_n\}$ and $\{w'_1, \dots, w'_m\}$ is

$$A' \;=\; T^{-1} \cdot A \cdot S \,.$$

Particular case: *The following particular case is of great importance for us.*

Assume that $V = W$ and $v_1 = w_1, \dots, v_n = w_n$. Let A (resp. A') be the matrix of a linear map $\alpha : V \to V$ with respect to the basis $\{v_1, \dots, v_n\}$ (resp. with respect to $\{w_1, \dots, w_n\}$). Let S be base change matrix. Then $A' = S^{-1} A S$.

Important remark. *Let $\alpha : V \to V$ be a linear map. Then the composition $\alpha \circ \alpha : V \to V$, $\ v \to \alpha(\alpha(v))$ is also linear. Furthermore, if A is the matrix of α with respect to some basis $\{v_1, \dots, v_n\}$ then the matrix of $\alpha \circ \alpha$ in the same basis is A^2. The similar remark applies to $\alpha \circ \alpha \circ \alpha$ and etc.*

Examples.

(i) The vector space F^n has the standard basis $e_1, \ldots, e_i$, where

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

with the "1" is in the ith row. Let $A = (a_{ij})$ be a $m \times n$-matrix. This matrix gives rise to an F-linear map

$$\alpha_A \,:\, F^n \longrightarrow F^m\,,\ v \longmapsto A \cdot v\,,$$

whose matrix with respect to the standard bases is precisely A. As we saw above choosing different bases on F^n and F^m will lead to another matrix.

For instance, let $n = m$ and $v_1, \ldots, v_n$, where $v_i = \begin{pmatrix} s_{1i} \\ \vdots \\ s_{ni} \end{pmatrix}$, be another basis of F^n. Then α_A has with respect to this basis the matrix

$$S^{-1} \cdot A \cdot S\,,$$

where $S = (s_{ij})$.

(ii) Let $V = \mathbb{C}$ be considered as $\mathbb{R}$-vector space, and $i = \sqrt{-1}$. The map $\ell_\iota : \mathbb{C} \longrightarrow \mathbb{C}$, $z \mapsto \iota \cdot z$ is $\mathbb{R}$- and also $\mathbb{C}$-linear. The complex numbers $\{1, \iota\}$ is a basis of the $\mathbb{R}$-vector space $\mathbb{C}$. We compute the matrix of ℓ_ι with respect to this basis. We have

$$\ell_\iota(1) \,=\, \iota \,=\, 0 \cdot 1 + 1 \cdot \iota$$

and

$$\ell_\iota(\iota) \,=\, -1 \,=\, (-1) \cdot 1 + 0 \cdot \iota\,,$$

and so the associated matrix is:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Exercise. What is the matrix with respect to the basis $\{1 + \iota,\, 1 - \iota\}$, and what are the base change matrices?

Subspaces invariant under a linear map.

We start with an example. Let

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 7 \\ 0 & 0 & 2 \end{pmatrix} \in \mathrm{M}_{3\times 3}(\mathbb{R}) .$$

The matrix A is of the form $\begin{pmatrix} B & C \\ \mathrm{O}_{1\times 2} & D \end{pmatrix}$, where B is a 2×2-matrix, C is a 2×1-matrix and D is a 1×1-matrix. Let $\alpha_A : \mathbb{R}^3 \to \mathbb{R}^3$, $v \to Av$ be the linear map corresponding to A.

We have $A \cdot e_1 = e_1 + 2e_2$ and $A \cdot e_2 = e_2$. Hence for the vector subspace $U \subset \mathbb{R}^3$ spanned by e_1, e_2 we have $\alpha_A(U) \subset U$. Such a subspace is called A- or α_A-invariant. We give a general definition.

Definition. *Let V be an F-vector space and $\alpha : V \longrightarrow V$ an F-linear map. We say that a subspace W of V is α-invariant if $\alpha(w) \in W$ for all $w \in W$, or in other symbols, if*

$$\alpha(W) := \big\{\alpha(w) \,|\, w \in W\big\} \subseteq W ,$$

or in words, if α maps W into W.

Remark. If $V = F^n$ and $\alpha = \alpha_A : v \mapsto A \cdot v$ for some $n \times n$-matrix A we say also that W is *A-invariant*.

Examples. Let $\alpha : V \longrightarrow V$ be an F-linear map. We denote $\alpha^n = \alpha \circ \alpha \ldots \circ \alpha$ (n-times), e.g. $\alpha^2(v) = \alpha(\alpha(v))$, $\alpha^3(v) = \alpha(\alpha(\alpha(v)))$ and so on.

(i) The kernel and the image of α is α-invariant.

(ii) Let A be a $n \times n$-matrix with eigenvalue $\lambda \in F$. The eigenspace of λ:

$$E_\lambda = \{v \in F^n \,|\, A \cdot v = \lambda \cdot v\}$$

is an A-invariant subspace of F^n. In fact, if $v \in E_\lambda$ then $A \cdot (A \cdot v) = A \cdot (\lambda v) = \lambda(A \cdot v)$ and so $A \cdot v$ is also in E_λ.

More generally, if $\alpha : V \longrightarrow V$ is an F-linear map then the eigenspace

$$E_\lambda = \{v \in V \,|\, \alpha(v) = \lambda \cdot v\}$$

of every eigenvalue λ of α is α-invariant. (Recall here that an eigenvalue of α is a zero of the characteristic polynomial of the matrix of α with respect to some basis of V. Recall also that the characteristic polynomial is independent of the choice of a basis.)

(iii) Let

$$W \;:=\; \left\{ v \in V \mid \alpha^n(v) = 0 \quad \text{for some } n \geq 1 \right\}.$$

This is a subspace of V since if $v, w \in W$, say $\alpha^m(v) = \alpha^n(w) = 0$ for some $m, n \geq 1$ then $\alpha^{m+n}(\lambda v + \mu w) = 0$ by linearity for all $\lambda, \mu \in F$. Moreover W is α-invariant since if $\alpha^n(v) = 0$ then $\alpha^{n-1}(\alpha(v)) = \alpha^n(v) = 0$.

(iv) Let $\alpha : V \longrightarrow V$ and λ be as in (ii) above. The subspace E_λ is equal to the kernel of $\alpha - \lambda\,\mathrm{id}_V$, where $\mathrm{id}_V : V \longrightarrow V$, $v \mapsto v$, is the identity mapping. Let

$$K_\lambda \;:=\; \left\{ v \in V \mid (\alpha - \lambda \cdot \mathrm{id}_V)^n(v) = 0 \quad \text{for some } n \geq 1 \right\},$$

the so called generalized eigenspace for the eigenvalue λ (this space plays an important role later and will be defined then again). As in (iii) above this is a subspace of V, and α-invariant since

$$\alpha \circ (\alpha - \lambda \cdot \mathrm{id}_V)^n \;=\; (\alpha - \lambda \cdot \mathrm{id}_V)^n \circ \alpha\,,$$

and so if $(\alpha - \lambda \cdot \mathrm{id}_V)^n(v) = 0$ then

$$(\alpha - \lambda \cdot \mathrm{id}_V)^n(\alpha(v)) \;=\; \alpha((\alpha - \lambda \cdot \mathrm{id}_V)^n(v)) \;=\; \alpha(0) \;=\; 0\,.$$