Linear Analysis
Lecture 3

Elementary Matrices

$E_{rs} \in \mathbb{F}^{m \times n}$ have a 1 is the (r, s)-entry and 0 elsewhere, e.g. in $\mathbb{F}^{4 \times 5}$

$$E_{24} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Note $C_f = E_{n1} + S_n$.

The elementary matrices form the *standard basis* for $\mathbb{F}^{m \times n}$.

Multiplication by Elementary Matrices

Left multiplication of $T \in \mathbb{F}^{m \times n}$ by $E_{rs} \in \mathbb{F}^{m \times m}$ moves the sth row of T to the rth row and zeros out all other elements.

That is, the elements of the matrix $E_{rs}T$ are all zero except for those in the rth row which is just the sth row of T.

Right multiplication of $T \in \mathbb{F}^{m \times n}$ by $E_{rs} \in \mathbb{F}^{n \times n}$ moves the rth column of T to the sth column and zeros out all other elements.

That is, the elements of the matrix TE_{rs} are all zero except for those in the sth column which is just the rth column of T.

$$(I + \alpha E_{rs})$$

Let $T \in \mathbb{F}^{n \times n}$. Left multiplication of T by the matrix $(I + \alpha E_{rs})$ adds α times the sth row of T to the rth row of T.

This is one of the elementary row operation used in Gaussian elimination.

Note that $E_{rs}^2 = 0$ whenever $r \neq s$, and so

$$(I - \alpha E_{rs})(I + \alpha E_{rs}) = I - \alpha E_{rs} + \alpha E_{rs} - \alpha E_{rs}^2 = I.$$

That is,

$$(I + \alpha E_{rs})^{-1} = (I - \alpha E_{rs}),$$

which makes sense since the inverse of adding α times the sth row to the rth row is to subtract it.

More Examples of Linear Transformations

Construct $G : \mathbb{F}^\infty \to \mathbb{F}^\infty$ by matrix multiplication. Let

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots \\ t_{21} & \ddots & \\ \vdots & & \end{pmatrix}$$

be an infinite matrix, each row having only finitely many nonzeros.

$$Tx \qquad \text{for} \qquad x = \begin{pmatrix} x_1 \\ \vdots \end{pmatrix} \quad \in \; \mathbb{F}^\infty$$

is well-defined since each entry in Tx is a finite sum.
Clearly, T is a linear transformation from $\mathbb{F}^\infty$ to itself.
Shift operators

$$(x_1, x_2, \ldots)^T \mapsto (0, x_1, x_2, \ldots)^T$$

and

$$(x_1, x_2, \ldots)^T \mapsto (x_2, x_3, \ldots)^T$$

are of this form.

More Linear transformations on Function Spaces

(a) Let $k \in C([c,d] \times [a,b])$ where $[a,b],[c,d]$ are closed bounded intervals. Define $L : C[a,b] \to C[c,d]$ by

$$L(u)(x) = \int_a^b k(x,y)u(y)dy.$$

L is called an **integral operator** with **kernel** $k(x,y)$.

(b) Let $m \in C[a,b]$. Then

$$L(u)(x) = m(x)u(x)$$

defines a **multiplier operator** L on $C[a,b]$.

(c) Let $g : [c,d] \to [a,b]$. Then

$$L(u)(x) = u(g(x))$$

defines a **composition operator** $L : C[a,b] \to C[c,d]$.

(d) $u \mapsto u'$ defines a **differential operator** $L : C^1[a,b] \to C[a,b]$.

Matrices and Basis Transformations

Let the vector space V and W over $\mathbb{F}$ have bases $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ and $\mathcal{B}_2 = \{w_1, \ldots, w_m\}$, resp.ly. Let $L : V \to W$ be linear.
For $1 \le j \le n$, write $Lv_j = \sum_{i=1}^{m} t_{ij} w_i$.
We call

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & & \vdots \\ t_{m1} & \cdots & t_{mn} \end{pmatrix} \in \mathbb{F}^{m \times n}$$

the matrix of L w.r.t. the basis $\mathcal{B}_1$ and $\mathcal{B}_2$.
Formally we have

$$L(v_1, \ldots, v_n) = (w_1, \ldots, w_m)\, T$$

Let $v \in V$ and $Lv \in W$ have $\mathcal{B}_1$ and $\mathcal{B}_2$ coordinates

$$v = (v_1, \ldots, v_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad Lv = (w_1, \ldots, w_n) \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

respectively. Then

$$(w_1, \ldots, w_n) y = Lv = L(v_1, \ldots, v_n) x = (w_1, \ldots, w_m)\, Tx \; .$$

That is, $y = Tx$.

Matrices and Basis Transformations

Let $\mathcal{B}_1' = \{v_1', \ldots, v_n'\}$ and $\mathcal{B}_2' = \{w_1', \ldots, w_m'\}$ be different basis for V and W with change-of-bases matrix $A \in \mathbb{F}^{n\times n}$ and $B \in \mathbb{F}^{m\times m}$, resp.ly. Then

$$(v_1' \cdots v_n') = (v_1 \cdots v_n)A \quad \text{and} \quad (w_1' \cdots w_m') = (w_1 \cdots w_m)B.$$

Consequently,

$$\begin{aligned} L(v_1' \cdots v_n') &= L(v_1 \cdots v_n)A \\ &= (w_1 \cdots w_n)\,TA \\ &= (w_1' \cdots w_m')B^{-1}\,TA\ . \end{aligned}$$

Therefore, the matrix of L in the new bases is $B^{-1}\,TA$.
If $W = V,\ \mathcal{B}_2 = \mathcal{B}_1,$ and $\mathcal{B}_2' = \mathcal{B}_1',$ then $B = A$, so the matrix of L in the new basis is $A^{-1}\,TA$.

Similarity Transformations

$A^{-1}TA$ said to be **similar** to T.

The transformation $T \mapsto A^{-1}TA$ is said to be a **similarity transformation** of A.

A similarity transformation of a matrix corresponds to the representation of the same linear transformation with respect to different bases.

Matrices vs Linear Transformations

Linear transformations can be studied abstractly or in terms of matrix representations.
For $L : V \to W$,

$$\text{the range} \quad \mathcal{R}(L),$$
$$\text{null space} \quad \mathcal{N}(L),$$
$$\operatorname{rank}(L) = \dim(\mathcal{R}(L)),$$
$$\text{etc.},$$

can be defined directly in terms of L, or in terms of matrix representations.

If $T \in \mathbb{F}^{n \times n}$ is the matrix of $L : V \to V$ in some basis, it is easiest to define

$$\det L = \det T \quad \text{and} \quad \operatorname{tr} L = \operatorname{tr} T.$$

Since

$$\det(A^{-1}TA) = \det T \quad \text{and} \quad \operatorname{tr}(A^{-1}TA) = \operatorname{tr}(T),$$

these are independent of the choice of basis.

Vector Spaces of Transformations

Let V, W be vector spaces.
If $L_1 : V \to W$, $L_2 : V \to W$ are linear, define

$$\alpha_1 L_1 + \alpha_2 L_2 : V \to W \quad \text{for } \alpha_1, \alpha_2 \in \mathbb{F} \text{ by}$$
$$(\alpha_1 L_1 + \alpha_2 L_2)v = \alpha_1 L_1(v) + \alpha_2 L_2(v) \; .$$

Therefore, the space of all linear transformations from V to W is naturally a vector space over $\mathbb{F}$ which we denote as $\mathcal{L}(V, W)$, or simply $\mathcal{L}(V)$ if $V = W$.

If V, W are finite-dimensional, we denote this vector space by $\mathcal{B}(V, W)$ so that

$$\mathcal{L}(V, W) = \mathcal{B}(V, W) \quad \text{and} \quad \mathcal{L}(V) = \mathcal{B}(V).$$

In the infinite-dimensional case, we use the notation $\mathcal{B}(V, W)$ to mean all *bounded* linear transformations (to be defined) from V to W with respect to norms on V, W.

$\mathcal{L}(V)$ as a Linear Algebra

A *linear algebra* over the field $\mathbb{F}$ is a vector space $\mathcal{L}(V)$ over $\mathbb{F}$ with an additional operation called *multiplication*.

The multiplication in a linear algebra associates with each pair of vectors $S, T \in \mathcal{L}$ a vector ST in $\mathcal{L}$ called the *product* of S and T.

Multiplication of vectors in a linear algebra satisfies

1. Associativity: $\forall\ R, S, T \in \mathcal{L}(V)\ R(ST) = (RS)T$
2. Distributivity wrt vector addition: $\forall\ R, S, T \in \mathcal{L}(V)$

$$R(S+T) = RS + RT \quad \text{and} \quad (R+S)T = RT + ST$$

3. Distributivity of scalar multiplication: $\forall\ \alpha \in \mathbb{F},\ S, T \in \mathcal{L}$

$$\alpha(ST) = (\alpha S)T = S(\alpha T)\ .$$

$\mathcal{L}(V)$ as a Linear Algebra

If there exists $I \in \mathcal{L}(V)$ such that

$$IS = S \qquad \forall\, S \in \mathcal{L},$$

then $\mathcal{L}(V)$ is called a linear algebra with identity, and I is called the identity.

The algebra $\mathcal{L}$ is called commutative if

$$ST = TS \qquad \forall\, S, T \in \mathcal{L}.$$

Fact For any vector space V, the vector space $\mathcal{L}(V)$ is a linear algebra with identity where multiplication on $\mathcal{L}(V)$ is defined as composition of two linear transformations of V to itself, and the identity is the identity transformation.

$\mathcal{L}(V)$ is not commutative if $\dim(V) > 1$.

Projections

V a vector space. Subspaces W_1 and W_2 are said to be *complementary* if

$$V = W_1 \oplus W_2 \ .$$

In this case, every $v \in V$ has a unique representation

$$v = w_1 + w_2 \quad \text{with} \quad w_1 \in W_1,\ w_2 \in W_2.$$

Using this decomposition, we define mappings

$$P_1 : V \to W_1 \quad \text{and} \quad P_2 : V \to W_2$$

by

$$P_1 v = w_1 \quad \text{and} \quad P_2 v = w_2 \ .$$

Each P_i is a mapping from V to itself.
P_1 is called the projection onto W_1 along W_2.
P_1 is determined by *both* W_1 and W_2.
P_i is linear: $v_1 = w_{11} + w_{12}$, $v_2 = w_{21} + w_{22}$, then

$$\begin{aligned} P_1(\alpha v_1 + \beta v_2) &= P_1((\alpha w_{11} + \beta w_{21}) + (\alpha w_{21} + \beta w_{22})) \\ &= \alpha w_{11} + \beta w_{21} = \alpha P_1(v_1) + \beta P_1(v_2) \ . \end{aligned}$$

Projections

It follows that

$$P_1^2 = P_1, \quad P_2^2 = P_2, \quad P_1 + P_2 = I, \quad \text{and} \quad P_1 P_2 = P_2 P_1 = 0.$$

An element q of an algebra is call **idempotent** if $q^2 = q$.

If $P : V \to V$ is a linear transformation and P is idempotent, then

P is the projection onto $\mathcal{R}(P)$ along $\mathcal{N}(P)$.

We extend this to $V = W_1 \oplus \cdots \oplus W_m$ for subspaces W_i.

Define projections $P_i : V \to W_i$ in the obvious way:

P_i the projection onto W_i along

$$W_1 \oplus \cdots \oplus W_{i-1} \oplus W_{i+1} \oplus \cdots \oplus W_m \ .$$

Then

$$P_i^2 = P_i \quad \text{for} \quad 1 \leq i \leq m, \qquad P_i P_j = P_j P_i = 0 \text{ for } i \neq j,$$

and

$$P_1 + \cdots + P_m = I.$$

Projections

If V is finite dimensional, we say that a basis

$$\{w_1, \dots, w_p, u_1, \dots, u_q\} \quad \text{for} \quad V = W_1 \oplus W_2$$

is **adapted to the decomposition** $W_1 \oplus W_2$ if

$$\{w_1, \dots, w_p\} \quad \text{is a basis for} \quad W_1$$

and

$$\{u_1, \dots, u_q\} \quad \text{is a basis for} \quad W_2.$$

Wrt such a basis, the matrix representations of P_1 and P_2 are

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix},$$

where the block structure is

$$\begin{bmatrix} p \times p & p \times q \\ q \times p & q \times q \end{bmatrix}, \quad \text{abbreviated} \quad \begin{array}{c} \\ p \\ q \end{array} \begin{array}{c} \begin{array}{cc} p & q \end{array} \\ \begin{bmatrix} * & * \\ * & * \end{bmatrix} \end{array}.$$

Invariant Subspaces

A subspace $W \subset V$ is *invariant* under $L : V \to V$ if $L(W) \subset W$.
Assume V has a basis $\{w_1, \dots, w_p, u_1, \dots, u_q\}$, where $\mathcal{B} = \{w_1, \dots, w_p\}$ is a basis for W. Then W is invariant under L iff the matrix of L in this basis is of the form

$$\begin{array}{cc} & \begin{array}{cc} p & q \end{array} \\ \begin{array}{c} p \\ q \end{array} & \begin{bmatrix} * & * \\ 0 & * \end{bmatrix}, \end{array} \quad \text{i.e., block upper-triangular.}$$

We say that $L : V \to V$ preserves the decomposition

$$V = W_1 \oplus \cdots \oplus W_m$$

if each W_i is invariant under L. In this case, L defines linear trans.'s $L_i : W_i \to W_i,\ 1 \leq i \leq m$, and write $L = L_1 \oplus \cdots \oplus L_m$.
L preserves the decomposition iff the matrix T of L with respect to an adapted basis is of block diagonal form

$$T = \begin{bmatrix} T_1 & & & 0 \\ & T_2 & & \\ & & \ddots & \\ 0 & & & T_m \end{bmatrix},$$

where the T_i's are the matrices of the L_i's in the bases of the W_i's.

Nilpotents

A linear transformation $L : V \to V$ is called **nilpotent** if

$$L^r = 0 \quad \text{for some } r > 0.$$

Shifts revisited:
The proto-typical example of a nilpotent operator on $\mathbb{F}^n$ is the shift operator S:

$$Se_i = e_{i-1} \quad (i \geq 2), \quad Se_1 = 0\ .$$

The matrix representation for S^m is

$$S^m = (S_n)^m = \begin{bmatrix} 0 & \cdots & 1 & \dots & 0 \\ & \ddots & & \ddots & \vdots \\ & & \ddots & & 1 \\ & & & \ddots & \vdots \\ 0 & & & & 0 \end{bmatrix} \quad \begin{matrix} \longleftarrow (1, m+1) \text{ element} \\ \longleftarrow (n-m, n) \text{ element.} \\ \\ \\ \\ \end{matrix}$$

Note, however that the shift operator on $\mathbb{F}^\infty$:

$$Se_i = e_{i-1} \quad (i \geq 2), \quad Se_1 = 0,$$

is **not** nilpotent.

Structure of Nilpotent Operators in Finite Dimensions

$L : V \to V$ is nilpotent and V finite dimentional.

Fact: *There is a basis of V in which L is the direct sum of shift operators.*

This decomposition results from a direct sum decomposition of V.
The decomposition is built on the structure of the subspaces $\mathcal{R}(L^k)$ and the corresponding the structure of the subspaces $\mathcal{N}(L^k)$.

This is a key step in showing that every matrix is similar to a matrix in Jordan form.

We look at the first step in the proof of this decomposition.

Structure of Nilpotent Operators in Finite Dimensions

L nilpotent $\Longrightarrow L^r = 0 \neq L^{r-1}$ for some r.

Let $v_1, \ldots, v_{\ell_1}$ be a basis for $\mathcal{R}(L^{r-1})$.

Choose $w_i \in V$ for which $v_i = L^{r-1} w_i$ $(1 \leq i \leq \ell_1)$.

Then

$$V = \mathcal{N}(L^r) = \mathcal{N}(L^{r-1}) \oplus \mathrm{Span}\{w_1, \ldots, w_{\ell_1}\}. \quad (1)$$

We claim that the set

$$\begin{aligned} \mathcal{S}_1 \quad = \quad \{ \; & L^{r-1} w_1, L^{r-2} w_1, \ldots, w_1, \\ & L^{r-1} w_2, L^{r-2} w_2, \ldots, w_2, \\ & \ldots, \\ & L^{r-1} w_{\ell_1}, L^{r-2} w_{\ell_1}, \ldots, w_{\ell_1} \; \} \end{aligned}$$

is linearly independent.

Structure of Nilpotent Operators in Finite Dimensions

Suppose

$$0 = \sum_{i=1}^{\ell_1} \sum_{k=0}^{r-1} c_{ik} L^k w_i \ .$$

Apply L^{r-1} to obtain

$$0 = \sum_{i=1}^{\ell_1} c_{i0} L^{r-1} w_i = \sum_{i=1}^{\ell_1} c_{i0} v_i,$$

The linear independence of the v_i's gives

$$c_{i0} = 0 \text{ for } 1 \leq i \leq \ell_1.$$

Now apply L^{r-2} to the double sum to obtain

$$0 = \sum_{i=1}^{\ell_1} c_{i1} L^{r-1} w_i = \sum_{i=1}^{\ell_1} c_{i1} v_i,$$

so $c_{i1} = 0$ for $1 \leq i \leq \ell_1$.
Successively applying lower powers of L shows that all $c_{ik} = 0$.

For $1 \leq i \leq \ell_1$,

$$\mathrm{Span}\{L^{r-1}w_i, L^{r-2}w_i, \ldots, w_i\}$$

is invariant under L, and L acts by shifting these vectors.
It follows that on $\mathrm{Span}(\mathcal{S}_1)$, L is the direct sum of ℓ_1 copies of the $(r \times r)$ shift S_r, and in the basis

$$\begin{aligned} \mathcal{S}_1 \quad = \quad \{ \; & L^{r-1}w_1, L^{r-2}w_1, \ldots, w_1, \\ & L^{r-1}w_2, L^{r-2}w_2, \ldots, w_2, \\ & \ldots, \\ & L^{r-1}w_{\ell_1}, L^{r-2}w_{\ell_1}, \ldots, w_{\ell_1} \; \} \end{aligned}$$

in the subspace $\mathrm{Span}(\mathcal{S}_1)$, L has the matrix

$$\begin{bmatrix} S_r & & 0 \\ & \ddots & \\ 0 & & S_r \end{bmatrix}.$$

In general, $\mathrm{Span}(\mathcal{S}_1)$ need not be all of V, so we aren't done.
Start this process again with $\mathcal{R}(L^{r-2})$, but with much greater care.

Consequences for Nilpotent Operators

Facts about nilpotents follow from this normal form.

For example, if $\dim V = n$ and $L : V \to V$ is nilpotent, then

(i) $L^n = 0$

(ii) $\operatorname{tr} L = 0$

(iii) $\det L = 0$

(iv) $\det (I + L) = 1$

(v) for any $\lambda \in \mathbb{F}$, $\det (\lambda I - L) = \lambda^n$