Linear Algebra Short Course

Lecture 2

Matthew J. Holland

matthew-h@is.naist.jp

Mathematical Informatics Lab
Graduate School of Information Science, NAIST

Some useful references

- Introduction to linear maps: Axler (1997, Ch. 3)
- Metric space of linear maps: Rudin (1976, Ch. 9)
- Excellent review of matrix basics: Horn and Johnson (1985, Ch. 0)
- Very accessible matrix algebra; basic identities, inequalities: Magnus and Neudecker (1999, Ch. 1–3,11)
- Invariant quantities: Axler (1997, Ch. 10) (note high dependency on previous chapters)

Lecture contents

1. **Linear transformations and their classes**
2. **Transformations and space structure**
3. **Matrices and their role in the theory**

Lecture contents

1. Linear transformations and their classes
2. Transformations and space structure
3. Matrices and their role in the theory

Linearity: from sets to functions

The “stage” for our current theory is vector spaces U, V, W with common field $\mathbb{F}$, assumed $\mathbb{R}$ or $\mathbb{C}$.

Our focus shifts from *sets* with a linearity property to *functions* with a linearity property.

Defn. We call $T : U \to W$ a **linear transformation** (or **map**) when $\forall\, u, u' \in U, \alpha \in \mathbb{F}$,

$$T(u + u') = T(u) + T(u')$$
$$T(\alpha u) = \alpha T(u)$$

(*) The naming is natural; T maps any linear combination of say $u_1, \ldots, u_m \in U$ to a linear combination of their maps $T(u_1), \ldots, T(u_m)$.

Linearity: from sets to functions

Some additional notation:

Denote by $\mathcal{L}(U, W)$ the set of all linear maps from U to W,

$$\mathcal{L}(U, W) := \{T : U \to W; T \text{ is linear}\}.$$

When $T \in \mathcal{L}(U, U)$, call T a **linear operator** on U.
Denote by $\mathcal{L}(U) := \mathcal{L}(U, U)$.

Linear operators are without question the key focus of LA.

Linear maps and bases

The bases of domain/co-domain of linear maps plays a key role.
Let $B_U = \{u_1, \ldots, u_m\}$ be a basis of U.

Example. (*) Linear maps on U are completely determined by where they map the vectors of B_U. That is, for linear maps $S, T \in \mathcal{L}(U, W)$,

$$S(u_i) = T(u_i), i = 1, \ldots, m \iff S = T.$$

Example. (*) Similarly, given arbitrary m vectors $w_1, \ldots, w_m \in W$, the *only* linear map $T \in \mathcal{L}(U, W)$ which satisfies $T(u_i) = w_i, i = 1, \ldots, m$ is that defined

$$T(u) := \alpha_1 w_1 + \cdots + \alpha_m w_m, \ \forall u \in U$$

where $u = \alpha_1 u_1 + \cdots + \alpha_m u_m$.

Various linear maps

(*) For $A \in \mathbb{R}^{m\times n}$, the map $S(x) := Ax$ is $S \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

(*) If $\mathcal{P}(\mathbb{R})$ is set of polynomials on $\mathbb{R}$, note

$$T(p) := \int_a^b p(x)\,dx \text{ satisfies } T \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathbb{R})$$

$$T(p) := p''(\cdot) \text{ satisfies } T \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R}))$$

(*) Counter-example: for $A \in \mathbb{R}^{m\times n}$, the map defined $S(x) := Ax + m$ for $m \neq 0$ is not linear, i.e., $S \notin \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

(*) For $x = (x_1, \ldots, x_n) \in \mathbb{F}^n$, note $T(x) := (x_{\pi(1)}, \ldots, x_{\pi(n)})$, where π is an arbitrary permutation, is $T \in \mathcal{L}(\mathbb{F}^n)$.

(*) T defined $(Tp)(x) := \beta x^3 p(x)$ for fixed $\beta \in \mathbb{R}$, is $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$.

Various linear maps (more)

(*) All linear operators on dim-1 spaces are simply scalar multiplications.

(*) Additivity is not a superfluous requirement; find a map $T : \mathbb{R}^2 \to \mathbb{R}$ such that $T(\alpha x) = \alpha T(x)$ but $T \notin \mathcal{L}(\mathbb{R}^2, \mathbb{R})$.

(*) Extensions of linear maps. Let $U \subset V$ be a subspace, and $T \in \mathcal{L}(U, W)$. Construct a map $\bar{T} \in \mathcal{L}(V, W)$ such that $\bar{T}(u) = T(u), \forall\, u \in U$.

Classes of linear maps

Linear spaces come in many varying forms.
With standard algebraic operations, $\mathcal{L}(U,V)$ is yet another example.

Example. (*) If U, V are vector spaces on field $\mathbb{F}$, define operations for arbitrary $S, T \in \mathcal{L}(U,V)$ by

$$(\alpha T)(\cdot) := \alpha T(\cdot),\ \forall\, \alpha \in \mathbb{F}$$
$$(T+S)(\cdot) := T(\cdot) + S(\cdot)$$

Consider what the additive inverse/identity are, recalling in particular VM.5 from Lec 1, and show $\mathcal{L}(U,V)$ is a vector space on $\mathbb{F}$.

What is $\dim \mathcal{L}(U,V)$? This motivates some new tools.

(*) If $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, note that the "operator norm"

$$\|T\| := \sup_{\|x\|_2 \le 1} \|T(x)\|_2$$

is a valid norm on $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$.

“Products” via compositions

A quasi-multiplication operation is naturally defined between elements of $\mathcal{L}(U, V)$ and $\mathcal{L}(V, W)$.

Defn. For $T \in \mathcal{L}(U, V)$, $S \in \mathcal{L}(V, W)$, we define the **product** ST by the composition

$$(ST)(u) := S(T(u)), \; \forall\, u \in U.$$

(*) As one would hope, $ST \in \mathcal{L}(U, W)$.

(*) Extends naturally to general case of $m \geq 2$ multiplicands, i.e., where $T_1 \in \mathcal{L}(V_0, V_1), T_2 \in \mathcal{L}(V_1, V_2), \ldots, T_m \in \mathcal{L}(V_{m-1}, V_m)$.

(*) The product is *almost* like that seen on fields. Prove:

- Analogue of associativity of multiplication on fields (FM.3).
- Existence of multiplicative identity, i.e., there exists $I \in \mathcal{L}(V, W)$ s.t. $IT = T$ for all $T \in \mathcal{L}(U, V)$, and vice versa.
- But commutativity need not hold, i.e., ST need not equal TS.

Lecture contents

1. <u>**Linear transformations and their classes**</u>
2. **Transformations and space structure**
3. **Matrices and their role in the theory**

Lecture contents

1. Linear transformations and their classes
2. Transformations and space structure
3. Matrices and their role in the theory

Transformation-induced structure

$T \in \mathcal{L}(U,V)$ induces all sorts of interesting *structure* to U, V.

Defn. The **nullspace** (or **kernel**) and **range** (or **image**) of T are

$$\begin{aligned} \operatorname{null} T &:= \{u \in U : Tu = 0\} \\ \operatorname{range} T &:= T(U) := \{v \in V : Tu = v, u \in U\} \end{aligned}$$

The structure we promised is easily observed.

(*) Both $\operatorname{null} T$ and $\operatorname{range} T$ are subspaces of U and V.

(*) Let $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ be the derivative operation. What is $\operatorname{null} D$?

(*) Same D but now $D \in \mathcal{L}(\mathcal{P}_k(\mathbb{R}))$, where $\mathcal{P}_k(\mathbb{R})$ restricts the polynomials to order $k > 0$ or less. What is $\operatorname{range} D$?

Transformation-induced structure

The key structural theorem for $T \in \mathcal{L}(U, V)$ is as follows.

Thm. (**) Let U be $\dim U < \infty$. Then, $\dim \operatorname{range} T < \infty$ and

$$\dim U = \dim \operatorname{null} T + \dim \operatorname{range} T.$$

This is a huge generalization of the key points of G. Strang's "fundamental theorems."

Example. (*) Let $A \in \mathbb{R}^{m \times n}$. Define $T(x) := Ax, S(x) := A^T y$. Then note

$$\operatorname{range} T = \operatorname{col} A = \operatorname{row} A^T, \quad \operatorname{range} S = \operatorname{col} A^T = \operatorname{row} A$$

and of course the nullspaces coincide with the usual nullspace of the matrices. The rest is just preservation of rowspaces in reducing to row-echelon form. The "rank" is just $\operatorname{rank} A = \dim \operatorname{range} T$.

Transformation info encoded in subspaces

A review of basic terms.

Defn. We call a map $T : U \to V$ **injective** if

$$u \neq u' \implies T(u) \neq T(u'),$$

and **surjective** if $\operatorname{range} T = V$.

If both, we call T **bijective**, or say it is a **one-to-one** mapping from U **onto** V.

(*) If $T \in \mathcal{L}(U, V)$ is injective and $\{u_1, \ldots, u_k\} \subset U$ is independent, then $\{T(u_1), \ldots, T(u_n)\} \subset V$ is independent. What about if not injective?

(*) Similarly, if $[\{u_1, \ldots, u_k\}] = U$ and T is surjective, then $[\{T(u_1), \ldots, T(u_k)\}] = V$. What if not surjective?

Transformation info encoded in subspaces

The structural results furnish handy conditions for these properties.

Assume general $T \in \mathcal{L}(U, V)$.

(*) T injective $\iff$ $\operatorname{null} T = \{0\}$.
(*) Thus injectivity equivalent to $\dim U = \dim \operatorname{range} T$.
(*) If $\dim U > \dim V$, then T cannot be injective.
(*) If $\dim U < \dim V$, then T cannot be surjective.
(*) Thus, have $\exists$ surjective $T \in \mathcal{L}(U, V) \iff \dim V \leq \dim U$.

Let $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. Statements about generalized linear systems follow naturally from these results:

(*) In terms of m and n, what can we say about the existence and uniqueness of solutions to $T(x) = 0$ and $T(x) = b$, $x \in \mathbb{F}^n, b \in \mathbb{F}^m$?

Invertibility of linear maps

Defn. We say $T \in \mathcal{L}(U, V)$ is **invertible** if $\exists\, T^{-1} \in \mathcal{L}(V, U)$ such that

$$T^{-1}T = I \in \mathcal{L}(U)$$
$$TT^{-1} = I \in \mathcal{L}(V)$$

where I is the identity map on the respective spaces.
Note: we are requiring T^{-1} be linear.

(*) Justify the notation T^{-1}; show the inverse, if it exists, is unique.

(*) The following fact should be verified.

$$T \text{ is invertible} \iff T \text{ is bijective}$$

The key to $\Longleftarrow$ direction is proving the inverse is *linear*.

Basic isomorphism theorems

Defn. If exists $T \in \mathcal{L}(U, V)$, T invertible, then we say U and V are **isomorphic**.

(*) If U, V are isomorphic, then

$$\dim U < \infty \iff \dim V < \infty$$

(*) Let $\dim U, \dim V < \infty$. Then

$$U \text{ and } V \text{ isomorphic} \iff \dim U = \dim V.$$

This important basic fact says we can *always* find invertible linear maps between any finite-dim U, V of equal dimension.

Specializing to linear operators

Things often become easier when we focus on linear operators, namely $T \in \mathcal{L}(U)$.

(*) Assuming $\dim U < \infty$, the following are equivalent:

(1) T is invertible
(2) T is injective
(3) T is surjective

The finite-dim requirement is not vacuous:
(*) Define $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$ by $(Tp)(x) := 5x^3 p(x)$. Note injectivity need not imply surjectivity.

(*) For U on $\mathbb{F}$ with $\dim U < \infty$ and $S, T \in \mathcal{L}(U)$, we have:

$$ST \text{ invertible} \iff S, T \text{ both invertible}$$
$$ST = I \iff TS = I$$
$$T = \alpha I, \text{ some } \alpha \in \mathbb{F} \iff ST = TS, \forall S \in \mathcal{L}(U)$$

Lecture contents

1. Linear transformations and their classes
2. Transformations and space structure
3. Matrices and their role in the theory

Lecture contents

1. **Linear transformations and their classes**
2. **Transformations and space structure**
3. **Matrices and their role in the theory**

Matrices as arrays of field elements

Defn. In general, a $m \times n$ **matrix** B on field $\mathbb{F}$ is simply an array,

$$B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}, \quad b_{ij} \in \mathbb{F}$$

with addition/multiplication operations defined.

Some notation:

$[b_{ij}] := B$. Let b_i be ith row entries; $b_{(j)}$ are jth column entries.

Recall for $B, B' \in \mathbb{F}^{m \times n}$, $C \in \mathbb{F}^{n \times l}$, $x \in \mathbb{F}^n$, $\alpha \in \mathbb{F}$,

$$B + B' = [b_{ij} + b'_{ij}]$$

$$\alpha B = [\alpha b_{ij}]$$

$$Bx = x_1 b_{(1)} + \cdots + x_n b_{(n)} = (b_1^T x, \ldots, b_m^T x)$$

$$BC = \begin{bmatrix} Bc_{(1)} & \cdots & Bc_{(l)} \end{bmatrix} = \begin{bmatrix} b_1^T C \\ \vdots \\ b_m^T C \end{bmatrix}.$$

The many faces of matrices

Matrices are quite multifaceted; in particular, we're interested in:

- **Matrices as linear maps**
- **Matrices as representations of linear maps**

The first is easy.
Already showed $B \in \mathbb{F}^{m \times n}$ specifies a map in $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. Countless matrix identities and inequalities are well-known and very useful (Magnus and Neudecker, 1999).

The latter is more subtle.
The basic idea is that there exist equivalence classes of matrices unified by a unique "underlying linear map" whose characteristics specify properties of *all* the matrices in the equivalence class.

Matrix representations of abstract objects

Let $T \in \mathcal{L}(U, V)$, $\dim U, \dim V < \infty$, and fix bases $B_U := \{u_1, \ldots, u_n\}, B_V := \{v_1, \ldots, v_m\}$. Recalling

$$T(u_j) = a_{1j}v_1 + \cdots + a_{mj}v_m, \quad 1 \leq j \leq n$$

uniquely represents each $T(u_j) \in V$, the scalars a_{ij} completely specify T.

Defn. Given the above discussion, we define

$$M(T; B_U, B_V) := \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

called the **matrix representation** of T.
If $U = V$, denote $M(T; B_U) := M(T; B_U, B_U)$.

(*) For fixed bases, note map $T \mapsto M(T) \in \mathbb{F}^{m \times n}$ is a bijection.

Matrix representations of abstract objects

Let $T \in \mathcal{L}(U,V), S \in \mathcal{L}(V,W)$. Fix bases B_U, B_V, B_W.

(*) Natural properties hold; the representation of the product is the product of the representations:

$$M(ST; B_U, B_W) = M(S; B_V, B_W) M(T; B_U, B_V)$$

Things extend naturally to vectors. For $u \in U$, define

$$M(u; B_U) := \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

where $u = \alpha_1 u_1 + \cdots + \alpha_n u_n$ is its B_U expansion.
(*) Then handily, verify

$$M(T(u); B_V) = M(T; B_U, B_V) M(u; B_U).$$

Additional properties of $T \mapsto M(T)$

(*) First, note $\mathbb{F}^{m\times n}$ is a vector space. What is $\dim \mathbb{F}^{m\times n}$?

(*) Then, note for U, V on field $\mathbb{F}$, and M defined by $T \mapsto M(T; B_U, B_V)$ for fixed bases, we have linearity, i.e.,

$$M \in \mathcal{L}(\mathcal{L}(U,V), \mathbb{F}^{m\times n})$$

and furthermore M is invertible.

(*) Using this, prove

$$\dim \mathcal{L}(U,V) = \dim(U)\dim(V).$$

(*) For $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ and $M(T) = [c_{ij}] \in \mathbb{F}^{m\times n}$ wrt standard bases,

$$T(x) = M(T)x = x_1 c_{(1)} + \cdots + x_1 c_{(n)}, \quad \forall x \in \mathbb{F}^n.$$

Matrix representations of abstract objects

Why is this useful? Fixing bases, we may equivalently consider

$$T(u) = v \in V \quad \text{or} \quad M(T(u)) = M(T)M(u).$$

The former is abstract (u, v might be functions, etc.).
The latter is concrete (typically $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$).

This idea is central to linear algebra!

It says some U, V and U', V' can be very different, yet the transformations $\mathcal{L}(U, V)$ and $\mathcal{L}(U', V')$ are fundamentally linked.

This "link" is explicitly captured by matrix representations.

Links between genuinely distinct spaces

Example. (*) Consider $T \in \mathcal{L}(\mathcal{P}_m(\mathbb{C}), \mathcal{P}_{m+2}(\mathbb{C}))$ and $S \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^{m+2})$ defined for $\beta \in \mathbb{C}$ by

$$(Tp)(x) := \beta x^2 p(x),\ p \in \mathcal{P}_m(\mathbb{C})$$
$$S(u) := (0, 0, \beta u_1, \ldots, \beta u_m),\ u \in \mathbb{R}^m$$

With respect to the "standard bases" of each space, verify

$$M(T) = M(S) = \begin{bmatrix} 0 & 0 & & 0 \\ 0 & 0 & & \\ \beta & 0 & & \\ 0 & \beta & & \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \beta \end{bmatrix},$$

a $(m+2) \times m$ complex matrix.

Some of the key questions of LA

Our next natural questions touch some of the fundamental objectives of linear algebra.

Let $T \in \mathcal{L}(U, V)$, with associated matrices

$$A := M(T; B_U, B_V)$$
$$A' := M(T; B'_U, B'_V).$$

What information about T can we decode from A and A'?

Is this information consistent between A and A'?

Decoding information from matrix representations

Defn. Let V be a vector space on $\mathbb{F}$, with $\dim V = n$. Given $G \in \mathbb{F}^{n\times n}$, if there exist bases $B_1 = \{b_1, \ldots, b_n\}, B_2 = \{b'_1, \ldots, b'_n\}$ such that

$$G = M(I; B_1, B_2) = \begin{bmatrix} M(b_1; B_2) & \cdots & M(b_n; B_2), \end{bmatrix}$$

then we call G a **change-of-basis matrix** on V from B_1 to B_2.
We shall often denote $G_{1,2} := G$ in this case.

(*) Every invertible $A \in \mathbb{F}^{n\times n}$ is a change of basis matrix.

(*) Conversely, every change of basis matrix is invertible, easily using the fact $I = M(I^2; B, B) = M(I; B, B')M(I; B', B)$.

The above facts are very important. Now we look at nomenclature.

Decoding information from matrix representations

(*) First note importantly that if $G_{1,2}$ is a change of basis matrix on V from B_1 to B_2, then

$$G_{1,2}^{-1} = G_{2,1}.$$

(*) With this, one may readily confirm

$$M(T;B_1) = G_{2,1}M(T;B_2)G_{1,2}.$$

Defn. We call two square matrices $A, B \in \mathbb{F}^{n\times n}$ **similar**, denoted $A \sim B$, if there exists a COB matrix G such that

$$A = G^{-1}BG.$$

(*) Note similarity "$\sim$" is an equivalence relation (i.e., check symmetry, reflexivity, transitivity).

Decoding information from matrix representations

So, *in the special case* of operator $T \in \mathcal{L}(U)$, we have

$$M(T;B) \sim M(T;B')$$

for *any* bases B, B'.

(*) Thus, if we know $A = M(T;B_1)$, and some $\bar{A} \sim A$, then it is guaranteed there exists a basis B_2 s.t.

$$\bar{A} = M(T;B_2).$$

Hence the equivalence class of matrices similar to $M(T;B_1)$ can be considered the class of matrices with "underlying map" T.

Decoding information from matrix representations

It is well-known that similar matrices $A \sim A'$ have many "invariants," such as:

- $\det A = \det A'$
- $\operatorname{trace} A = \operatorname{trace} A'$
- A and A' share eigenvalues
- A and A' share a characteristic polynomial
- The same "canonical forms" (sparse, convenient forms)

These facts are very nice for decoding information about *two distinct matrices* (since knowing similarity is sufficient).

However, this tells us little intrinsic information about T itself!

Can we define invariants in terms of *just* T that coincide with the invariants of its matrix representations?

The answer is *yes*; this will be handled in Lecture 3 mainly.

Issues with more general arguments

All the key ideas we just briefly introduced related to linear operator $T \in \mathcal{L}(U)$, with square matrix $M(T; B)$. Which assumptions are critical?

The deepest results are for operators only. Thus $T \in \mathcal{L}(U, U) = \mathcal{L}(U)$.

Of course this implies for any bases B, B' of U that $M(T; B, B')$ is square, but that is not enough; we are interested in matrix representations $M(T; B, B)$ only (save for the COB matrix).

The reason for this is in the next example.

Issues with more general arguments

Example. (*) Let $T \in \mathcal{L}(U)$ with $n \times n$ matrix

$$M(T; B_1, B_2) = [a_{ij}],$$

and let A' be the same as A, save for the kth row, which is defined

$$a'_k := a_k + \lambda a_l,$$

that is, by an "elementary operation" on A. Find a basis B'_2 such that

$$A' = M(T; B_1, B'_2).$$

Properties such as trace and determinant need not be preserved over such operations, and extensions as above are infeasible.

(*) Rank is interesting; recall it is preserved across elementary operations (which need not preserve similarity). It is also preserved across similar matrices, i.e., $A \sim A' \implies \operatorname{rank} A = \operatorname{rank} A'$.

Lecture contents

1. Linear transformations and their classes
2. Transformations and space structure
3. **Matrices and their role in the theory**

Lecture contents

1. Linear transformations and their classes
2. Transformations and space structure
3. Matrices and their role in the theory

References

Axler, S. (1997). *Linear Algebra Done Right*. Springer, 2nd edition.

Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge University Press, 1st edition.

Magnus, J. R. and Neudecker, H. (1999). *Matrix differential calculus with applications in statistics and econometrics*. Wiley, 3rd edition.

Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition.