Lecture 1 — Linear Algebra — Fall 2019

Outline: review of basics, spectral theory, special matrices, perturbation theory, multilinear algebra and tensors

Useful Books:

- Introduction to matrix analysis — R. Bellman
- Linear Algebra — Hoffman and Kunze
- The theory of matrices I and II — Gantmacher
- Determinants — Muir
- Linear Algebra — Lax

Facts and Results from the elementary theory of matrices that you should know:

vector spaces V over a field $\mathcal{F}$, vectors, subpaces $W \subset V$:

usually over the reals $\mathbb{R}$ or the complex numbers $\mathbb{C}$,

but the <u>algebra</u> of most of the elementary theory, as opposed to the <u>analysis</u>, goes through for vector spaces over any field $\mathcal{F}$ e.g. the field of 2 elements $\{0,1\}$.

inner product (x,y), x and y in V: say $(V, (\cdot,\cdot))$ is an inner product space.

For example: $(x, y) = \sum_{i=1}^{n} x_i y_i$ over $\mathbb{R}$, $x = (x_1, \ldots, x_n) \in \mathbb{R}$

$y = (y_1, \ldots, y_n) \in \mathbb{R}$

$(x, y) = \sum_{i=1}^{n} \overline{x_i} y_i$ over $\mathbb{C}$,

$x = (x_1, \ldots, x_n) \in \mathbb{C}^n$

$y = (y_1, \ldots, y_n) \in \mathbb{C}^n$.

$\|x\| \equiv \sqrt{(x, x)}$ is a norm on V

$\|x\| = 1 \iff x$ is <u>normalized</u>

If $(x, y) = 0$, then x and y are orthogonal.

Independence of vectors: $c_1 x_1 + \cdots + c_n x_n = 0$
$\{x_1, \ldots, x_n\}$ in V for $c_i \in \mathcal{F}$,

$\Rightarrow c_1 = c_2 = \cdots = c_n = 0$

dependence of vectors: $c_1 x_1 + \cdots + c_n x_n = 0$
$\{x_1, \ldots, x_n\}$ in V with $c_i \neq 0$ for some i

Kronecker delta function: $\delta_{ik} = 1$ if $i = k$, $\delta_{ik} = 0$ if $i \neq k$.

Concept of a basis: $\{x_1, \ldots, x_n\}$ for V

Every vector space V has a basis

dimension of a vector space: $\dim V$

orthonormal basis: $\{x_1, \ldots, x_n\}$ is a basis with

$(x_i, x_j) = \delta_{ij}$ $\quad 1 \le i, j \le n$

Every vector space $(V, (\cdot, \cdot))$ has an orthonormal basis

matrices: $A = (a_{ij})_{1 \le i \le n, 1 \le j \le m}$ is an $n \times m$ matrix

matrix multiplication: $A = (a_{ij})$ $n \times m$, $B = (b_{ij})$, $m \times k$

$(AB)_{ij} = \sum_{\ell=1}^{m} a_{i\ell} b_{\ell j}$: AB is $n \times k$

transpose A^T of A : if $A = (a_{ij})$, $n \times m$
$A^T = (a_{ji})$, $m \times n$

adjoint A^* of A (in complex case) : $(A^*)_{ij} = (\bar{a}_{ji})$
$= \overline{A^T}$

row rank of a matrix A = maximal # of independent row vectors
column rank of a matrix A = maximal # of independent column vectors
Always have row rank = column rank $\equiv r(A)$

An $m \times n$ matrix $A = (a_{ij})$ induces a linear map from $\mathbb{R}^n$ (or $\mathbb{C}^n$) to $\mathbb{R}^m$ (or $\mathbb{C}^m$), via

$$(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j , \qquad 1 \le i \le m .$$

where $x = (x_1, \dots, x_n)$

Let A be a linear map from V_1 to V_2; written
$A \in \mathcal{L}(V_1, V_2)$

Then

null space of $A \equiv N(A) = \{ x \in V_1 : Ax = 0 \}$

range space of $A \equiv R(A) = \{ y \in V_2 : \exists x \in V_1$ such that $Ax = y \}$.

$N(A)$ and $R(A)$ are subspaces of V

Fundamental theorems for $A \in \mathcal{L}(V) \equiv \mathcal{L}(V, V)$:

(3.1) $\dim R(A) = r(A)$

(3.2) $\dim N(A) + \dim R(A) = n = \dim V$

(3.3) $(\text{Ran } A)^\perp = N(A^*)$ over $\mathbb{C}$
$= N(A^T)$ over $\mathbb{R}$

where

$$W^\perp = \{ y \in V : (y, x) = 0 \quad \forall x \in V \}$$

for any subspace $W \subset V$. $W^\perp$ is also a subspace of V

Note that (3.2) implies $A \in \mathcal{L}(V)$ is a surjection if and only if it is an injection.

Inverse A^{-1} of a square matrix : $A^{-1}A = AA^{-1} = I = (\delta_{ij})$ identity matrix. By the preceding comment note that

$A^{-1}A = I$ (resp $AA^{-1} = I$) $\Rightarrow$ $AA^{-1} = I$ (resp. $A^{-1}A = I$)

Let L be a linear map between vector spaces, $L \in \mathcal{L}(U, V)$

$$L : U \to V$$

Suppose $u_1, \dots, u_n$ is a basis for U and $v_1, \dots, v_m$ is a ~~c~~basis for V. Then ~~for~~ some $m \times n$ matrix

$(l_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$,

(4.1) $$L u_j = \sum_{i=1}^{m} l_{ij} v_i \quad , \quad 1 \le j \le n.$$

(l_{ij}) is called the <u>matrix associated</u> with L in the bases $\{u_j\}$ and $\{v_i\}$

<u>Exercise</u>: How does (l_{ij}) change if we change bases in U & V

<u>Important observation</u>: suppose $L \in \mathcal{L}(U,V)$ and $K \in \mathcal{L}(V,W)$.

Then $(KL)u \equiv K(Lu)$ defines by composition a linear map from U to W. Suppose $(u_1, \dots, u_n)$, $(v_1, \dots, v_m)$ and $(w_1, \dots, w_k)$ are bases for U, V and W respectively.

Let (l_{ij}) and (k_{ij}) be the matrices associated with L and K in these bases, respectively. Then

$$
\begin{aligned}
(KL)(u_i) &= K(Lu_i) \\
&= K \sum_{j=1}^{m} l_{ji} v_j \\
&= \sum_{j=1}^{m} l_{ji} \, Kv_j \\
&= \sum_{j=1}^{m} l_{ji} \sum_{r=1}^{k} k_{rj} w_r \\
&= \sum_{r=1}^{k} \left(\sum_{j=1}^{m} k_{rj} l_{ji} \right) w_r
\end{aligned}
$$

Thus the matrix associated with KL in the bases $\{u_i\}$ and $\{w_r\}$ is obtained by multiplying the matrix (k_{rj}) associated with K and the matrix (l_{ji}) associated with L using the <u>usual rules of matrix multiplication</u>. A different way of saying this is that matrix multiplication is defined in such a way as to make the map

$$L \to (l_{ij})$$

a homomorphism, i.e.

$$L + M \to (l_{ij}) + (m_{ij}) \qquad , L, M \in \mathcal{L}(U,V)$$

$$LK \to \underbrace{(l_{ij})(k_{ij})}_{\text{matrix multiplication}} \qquad , L \in \mathcal{L}(U,V)$$
$$K \in \mathcal{L}(V,W).$$

Systems of linear equations and elementary row operations: solving $Ax = b$ by Gaussian elimination

Here Ax is matrix multiplication of (a_{ij}) with the column vector $x = (x_1, \dots, x_n)^T$

elementary row operations: (1) add a multiple of one row to another

(2) interchange 2 rows

(3) multiply a row by a scalar

Exercise: Show that (1) (2) and (3) can be implemented by multiplying the matrix A on the left by suitable matrices E.

What happens if we multiply A on the right by such E's?

symmetric matrices : $A = A^T$

hermitian matrices : $A = A^*$

real matrices : $A = \bar{A}$

orthogonal matrices : $A^TA = I = AA^T$

unitary matrices : $A^*A = I = AA^*$

Thus $A^{-1} = A^T$ (resp. $A^{-1} = A^*$) in these cases

lower triangular matrics $A = (a_{ij})$: $a_{ij} = 0$ if $i < j$

upper triangular matrices $A = (a_{ij})$: $a_{ij} = 0$ if $i > j$

Note that if A is lower triangular (resp. upper triangular) and invertible, then A^{-1} is also lower triangular (resp. upper triangular).

$$\begin{array}{c} e_i = (0 \ldots 0 \; 1 \; 0 \ldots 0) \\ \uparrow_{i} \end{array}, \quad 1 \le i \le n$$ is the <u>standard basis</u>

for $\mathbb{R}^n$ or $\mathbb{C}^n$ regarded as a row space and

and e_i^T, $1 \le i \le n$, is the standard basis for $\mathbb{R}^n$ or

$\mathbb{C}^n$ regarded as a column vector space.

permutation matrix, diagonal matrix $D = (d_i \delta_{ij})$,

skew symmetric matrix $A^T = -A$

<u>Exercise</u> Use Gaussian elimination to show that all matrices A have a factorization $A = PLDU$ where P is a permutation matrix, L is lower, U is upper and D is

diagonal. Give a condition which ensures that $P = I$.

Determinant of an $n \times n$ matrix A:

$$|A| = \det A = \sum_{\sigma} (\operatorname{sgn} \sigma)\, a_{1\sigma_1} a_{2\sigma_2} \cdots a_{n\sigma_n}$$

Properties: (1) $\det AB = (\det A)(\det B)$

(2) interchanging 2 rows (or columns) changes the sign of the determinant

(3) $\det A$ is a multilinear function of the rows and columns of A, e.g. for any $1 \le i \le n$,

$$\begin{vmatrix} a_{11} & \cdots & \alpha a_{1i} + \beta b_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & \alpha a_{ni} + \beta b_{ni} & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{vmatrix}$$

$$+ \beta \begin{vmatrix} a_{11} & \cdots & b_{1i} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_{ni} & \cdots & a_{nn} \end{vmatrix}$$

(2) $\Rightarrow$: $\det A = 0$ if 2 rows (or colums) of A are equal

(2)(3) $\Rightarrow$: the determinant remains unchanged if a row (or column) is added to another row or column

(4) $\det A = \det A^T$

Exercise: Use (4) to show row rank = column rank

Exercise: If A is a skew symmetric matrix in an odd dimensional space, then $\det A = 0$

Equivalences: Suppose A is an $n \times n$ matrix. The following are equivalent

(1) $\det A \neq 0$ (2) $Ax = 0 \iff x = 0$

(3) $Ax = b$ is solvable (uniquely) for all b

(4) $r(A) = n$ (5) A^{-1} exists.

Exercise: Let V be the vector space of vectors $x = (x_1, x_2, x_3)$ where $x_i \in \mathbb{Z}_2 = \{0, 1\}$ with field $\mathbb{Z}_2$

Let $A \in \mathcal{L}(V)$, $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

Show that (1), ..., (5) all fail for A

Exercise: If A is a skew symmetric $n \times n$ matrix, show that $\dim (R(A))$ is even.

(Hint: A is one-to-one on $R(A)$)

Cofactors If $A = (a_{ij})$ is $n \times n$, then the complement of the entry a_{ij} is the $(n-1) \times (n-1)$ determinant obtained by deleting the ith row and jth column of A. The cofactor α_{ij} of a_{ij} is $(-1)^{i+j}$(complement of a_{ij})

e.g. for $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, then the complement of a_{23} is $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix}$

and $\alpha_{23} = (-1)^{2+3} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix}$

<u>Theorem</u>: Suppose A is $n \times n$ and $\det A \neq 0$.
Then

(10.1) $$(A^{-1})_{ij} = \frac{\alpha_{ji}}{\det A}$$

<u>Proof</u>: Recall Euler's Theorem for homogeneous functions $f: \mathbb{R}^n \to \mathbb{R}$,

$$f(\lambda x) = \lambda^p f(x) \qquad , \lambda > 0$$

for some p. For example, if $f(x) = \sqrt{x_1^2 + x_2^2}$; have

$f(\lambda x) = \lambda f(x)$ ie $p = 1$.

$$\text{Euler} \Rightarrow \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} x_i = p\, f(x)$$

(Prove this!).

Now clear as $\det A$ is linear with respect to each column of $\det A$, it is certainly homogeneous. Thus for any i,

$$\sum_{j=1}^n \frac{\partial \det A}{\partial a_{ji}} a_{ji} = \det A$$

But it is easy to see that

$$\frac{\partial \det A}{\partial a_{ji}} = \det \begin{vmatrix} a_{11} & a_{1i-1} & 0 & a_{1i+1} & a_{1n} \\ & & \vdots & & \\ a_{n1} & a_{ni-1} & 0 & a_{ni+1} & a_{nn} \end{vmatrix} = \alpha_{ji}$$

Thus we obtain the well-known expansion of a det in cofactors:

$$\sum_{j=1}^{n} \alpha_{ji}\, a_{ji} = \det A \tag{11.1}$$

For $k \neq i$ replace column i in A with column k, leaving all the remaining columns unchanged; call the new matrix $\tilde{A}$. We have from (11.1)

$$\sum_{j=1}^{n} \tilde{\alpha}_{ji}\, \tilde{a}_{ji} = \det \tilde{A} = 0$$

as $\tilde{A}$ has 2 equal columns. But $\tilde{a}_{ji} = a_{jk}$ and also $\tilde{\alpha}_{ji} = \alpha_{ji}$. Hence $\sum_{j=1}^{n} \alpha_{ji}\, a_{jk} = 0$. Thus

$$\frac{1}{\det A} \sum_{j=1}^{n} \alpha_{ji}\, a_{jk} = \delta_{ik}$$

which proves (10.1).

<u>Cramer's rule</u>: If $\det A \neq 0$ and $Ax = b$, then

$$x_i = \det \begin{pmatrix} a_{11} & a_{1i-1} & b_1 & a_{1i+1} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots a_{ni-1} & b_n & a_{ni+1} & & a_{nn} \end{pmatrix} \Big/ \det A \tag{11.1}$$

<u>Proof</u>: $x_i = (A^{-1}b)_i = \sum_{j=1}^{n} (A^{-1})_{ij}\, b_j = \frac{1}{\det A} \sum_{j=1}^{n} \alpha_{ji}\, b_j$

by (10.1). Now check that this is the same as expanding the determinant in (11.1) down the i^{th} column. $\square$

Example: Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ and solve $Ax = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

Show $\det A \neq 0$.

Then $x_3 = \det \begin{pmatrix} 1 & 2 & 1 \\ 4 & 5 & 0 \\ 7 & 8 & 0 \end{pmatrix} \Big/ \det A$.

Given vectors $v_1, \ldots, v_j \in V$, $\langle v_1, \ldots, v_k \rangle =$ span$(v_1, \ldots, v_j) = \{$all lin. combinations $c_1 v_1 + \cdots + c_j v_j$

Gram-Schmidt procedure: Given k independent vectors in an n dimensional space, $u_1, \ldots, u_k$, we can construct k orthonormal vectors $v_1, \ldots, v_k$ which span the same subspace as the u_i's, i.e.

$$\langle v_1, \ldots, v_k \rangle = \langle u_1, \ldots, u_k \rangle$$

Note: If $w_1, \ldots, w_j$ is a set of orthogonal, non-zero vectors, then $\{w_1, \ldots, w_j\}$ are indep

Step 1 As the u_i's are independent then certainly $u_1 \neq 0$:

set $v_1 = \pm u_1 / \|u_1\|$

Note $\|v_1\| = 1$

Step 2 Set $v_2 = \pm \dfrac{u_2 - (v_1, u_2) v_1}{\|u_2 - (v_1, u_2) v_1\|}$

Note (i) as u_2 and u_1, and hence u_2 and v_1, are independent, $u_2 - (v_1, u_2) v_1 \neq 0$

(ii) $\|v_2\| = 1$ and $(v_2, v_1) = 0$: in particular $\{v_1, v_2\}$ are indep

(iii) $\langle v_1, v_2 \rangle = \langle u_1, u_2 \rangle$

Step 3 Set $v_3 = \pm \dfrac{u_3 - (v_1, u_3) v_1 - (v_2, u_3) v_2}{\| u_3 - (v_1, u_3) v_1 - (v_2, u_3) v_2 \|}$

Note (i) as $v_1, v_2 \in \langle u_1, u_2 \rangle$, and as u_1, u_2, u_3 are independent, we have $u_3 - (v_1, u_3)v_1 - (v_2, u_3)u_2 \neq$

(ii) $\|v_3\| = 1$, $(v_1, v_3) = 0$, $(v_2, v_3) = 0$

(iii) $\langle v_1, v_2, v_3 \rangle = \langle u_1, u_2, u_3 \rangle$

etc. Continuing we construct $\{v_1, .., v_k\}$ with the desired properties

Let U be the $n \times k$ matrix with colums $u_1, .., u_k$ in that order, and let V be the $n \times k$ matrix with columns $v_1, .., v_k$, in that order.

Exercise Show that

(13.1) $$U = VR$$

where R is an upper triangular matrix with $R_{ii} \neq 0$

Choosing the $\pm$ signs in Step 1, Step 2, .. appropriately, we can assume $R_{ii} > 0$.

Now let L be an invertible $n \times n$ matrix. Then applying the Gram-Schmidt procedure to the column of L

starting from the left we obtain the factorization as above

(14.1) $L = QR$

As the columns of Q are orthonormal, Q is an orthogonal matrix $QQ^T = Q^TQ = I$ (or in the complex case, Q is unitary, $QQ^* = Q^*Q = I$); R is upper and we can assume $R_{ii} > 0$. (14.1) is called the QR factorization of L.

Exercise Show that the QR factorization of L, $\det L \neq 0$, is unique (we always assume $R_{ii} > 0$).

(14.1) Exercise: Suppose $u_1, \dots, u_j$ are orthog. and set $u = c_1 u_1 + \dots + c_j u_j$ for any $c_1, \dots, c_j$. Then $\|u\|^2 = |c_1|^2 \|u_1\|^2 + \dots + |c_j|^2 \|u_j\|^2$

Let $B = (b_{ij})$ be an $n \times n$ matrix.

Then

(14.2) $$|\det B| \leq \prod_{i=1}^{n} \left(\sum_{k=1}^{n} |b_{ik}|^2 \right)^{\frac{1}{2}}.$$

This is called Hadamard's Inequality

Proof: If $\det B = 0$, this is trivial. So assume $\det B \neq 0$, and let $B = QR$ be the QR factorization of B. Then for any i,

$$\begin{pmatrix} b_{1i} \\ \vdots \\ b_{ni} \end{pmatrix} = Q \begin{pmatrix} R_{1i} \\ \vdots \\ R_{ii} \\ 0 \\ 0 \end{pmatrix} = R_{1i} q_1 + \cdots + R_{ii} q_i$$

where q_j is the j^{th} column of Q. But then

by (14.1) $\sum_{j=1}^{n} |b_{ji}|^2 = |R_{1i}|^2 \|q_1\|^2 + \cdots + |R_{ii}|^2 \|q_i\|^2$

$$= |R_{1i}|^2 + \cdots + |R_{ii}|^2 \geq |R_{ii}|^2$$

as the q_j's are orthonormal.

As $|\det Q| = 1$, we have

$$|\det B| = |\det Q| \, |\det R|$$
$$= |\det R| = \prod_{i=1}^{n} |R_{ii}| \leq \prod_{i=1}^{n} \left(\sum_{j=1}^{n} |b_{ji}|^2 \right)$$

as desired.

<u>Exercise</u> Interpret Hadamard's inequality geometrically in terms of volumes.