Summary of Lecture II

Fundamental result: If V is generated by $v_1, \ldots, v_n$, and if $N > n$, then each set $w_1, \ldots, w_N$, w's $\in V$, is linearly dependent.

[Recall: A set S is linearly dependent if $\exists$ $s_1, \ldots, s_\ell$ distinct elements in S and $\alpha_1, \ldots, \alpha_\ell$, not all 0, $\ni$ $\sum_{i=1}^{\ell} \alpha_i s_i = 0$. Linearly independent means not linearly independent. Sometimes, the convention is added, and we do this, that a listed set $w_1, \ldots, w_\ell$ is dependent if it contains repetitions, since if a w occurs twice there is a nontrivial linear combination $w - w = 0$. This convention prevents us from having to say, e.g., in the Fundamental Result that the ws are distinct.]

We shall give two proofs of this Fund Result. The first proof involves a basic result about linear equations:

Theorem: A system of n homogeneous (linear) equations in N unknowns, $N > n$, has a nontrivial solution. (nontrivial means not all unknowns $= 0$).

Proof of the Theorem: Induction on the number of equations. It is "obvious" that one equation in two or more unknowns has a nontrivial solution: $a_1 x_1 + \ldots + a_N x_N = 0$, if all a_i's $\equiv 0$, has $x_1 = \ldots = x_N = 1$ as a solution, for example.

If some $a \neq 0$, say $a_1 \neq 0$ (wolog), then $x_2, \dots, x_N$ can be chosen arbitrary (and in particular nonzero) and x_1 can be taken to be $x_1 = -\frac{1}{a_1}(a_2 x_2 + \dots + a_N x_N)$.
For the induction step, look at the system

$$\begin{aligned} a_1^1 x_1 + \dots + a_N^1 x_N &= 0 \\ &\vdots \\ a_1^n x_1 + \dots + a_N^n x_N &= 0 . \end{aligned} \qquad (*)$$

The case of all a's $= 0$ is easy as before.
If not all a's $= 0$, then wolog (rearranging equations and relabelling the variables), we can assume $a_1^1 \neq 0$. The system $(*)$ has the same set of solutions as $(**)$ obtained by eliminating x_1 from all but the first equation:

$$\begin{aligned} a_1^1 x_1 + \dots + a_N^1 x_N &= 0 \\ 0 x_1 + \left(a_2^2 - \frac{a_1^2}{a_1^1} a_2^1\right) x_2 + \dots &= 0 \\ &\vdots \\ 0 x_1 + \left(a_2^n - \frac{a_1^n}{a_1^1} a_2\right) x_2 + \dots &= 0 \end{aligned} \qquad (**)$$

(usual elimination process).
The last $(n-1)$ equations in $N-1$ unknowns has by induction a nontrivial solution. Using

the first equation to get x_1 (possible since $a_1^1 \neq 0$) gives a nontrivial solution to $(*)$.

Example
$$\begin{aligned} 3x + 7y + z &= 0 \\ 2x + 6y - z &= 0 \end{aligned} \qquad (*)$$
has same solution set as
$$3x + 7y + z = 0$$
$$(2x + 6y - z) - \left(\tfrac{2}{3}\right)(3x + 7y + z) = 0$$
or
$$\begin{aligned} 3x + 7y + z &= 0 \\ 0x + \left(6 - \tfrac{14}{3}\right)y + \left(-1 - \tfrac{2}{3}\right)z &= 0 \end{aligned} \qquad (**)$$

If (y_0, z_0) is a nontrivial solution of the second equation, then

$\left(-\frac{1}{3}(7y_0 + z_0), y_0, z_0\right)$ is a nontrivial solution of $(*)$.

Now we turn to the proof (first proof) of the Fundamental Result:

Proof I: Write $\quad w_i = \sum_{j=1}^{n} \alpha_j^i v_j \qquad i = 1, \ldots N$

Then $\quad \sum_{i=1}^{N} x_i w_i = \sum_{i=1}^{N} \left(\sum_{j=1}^{n} x_i \alpha_j^i v_j \right)$
$$= \sum_{j=1}^{n} \left[\left(\sum_{i=1}^{N} x_i \alpha_j^i \right) \right] v_j$$

So $\quad \sum x_i w_i = 0$ if , for each $j = 1, \ldots n$,
$$\sum_{i=1}^{N} x_i \alpha_j^i = 0 .$$

This gives n equations in N unknowns $x_1, \ldots, x_N$. By the preliminary theorem, these n equations in N unknowns, $N > n$, have a nontrivial solution. Thus there exist $x_1, \ldots, x_N$ not all $= 0$ with

$$\sum_{j=1}^{N} x_j w_j = 0$$

and the w's are linearly dependent as required. $\square$

Proof 2 ("by legerdemain"): With $v_1, \ldots, v_n$ generating and $w_1, \ldots, w_N$ independent, we proceed as follows: For contradiction, assume $w_1 \cdots w_N$ independent, $N > n$.

The set $w_1\ v_1 \ldots v_n$ is dependent, since w_1 is a linear combination of the v's. Some v is involved in this linear combination. Choose the last (highest index) v that is involved (in some chosen linear combination) and remove it. Then

$$w_1\ v_1 \ldots \ldots v_n$$

some v removed

still generates V. Continue to get

$$w_2\ w_1\ v_1 \ldots \ldots v_n$$

two v's remo

that generates. Note that in this pr

there is always a v involved ,

w's by themselves are independ

are choosing the <u>last</u> item in

linear combination to remove. So eventually we get to where all v's are gone: and so $w_n \, w_{n-1} \, \ldots \, w_1$

generate V. But then
$$w_{n+1} \; w_n \; w_{n-1} \; \ldots \; w_1$$
is a dependent set since w_{n+1} = a linear combination of $w_n \, w_{n-1} \, \ldots \, w_1$.
(Note that $N > n \Rightarrow w_{n+1}$ exists!).
This is a contradiction since $w_1, \ldots, w_N$ was assumed to be an independent set. $\square$

Application:
Theorem (fundamental theorem on dimension):
If V is finite dimensional, then there is a finite set $v_1, \ldots, v_k$ of vectors in V such that this set generates V and is linearly independent, and all such sets contain the same number of elements.

Note:
Such a set is called a basis for V and the number k is called the dimension of V.

Proof. If $v_1, \ldots, v_n$ generate, then an independent generating set can be obtained by successively removing v's that are linear combinations of the remaining v's. (Details are an exercise).

If $v_1, \ldots, v_k$ and $w_1, \ldots, w_\ell$ are both independent generating sets then the Fundamental result from the beginning gives $\ell \leq k$ (because the w's are independent and the v's generate) and $k \leq \ell$ (because the v's are independent and the w's generate). So $k = \ell$. $\square$

There is another way to get a basis for a finite-dimensional vector space, which is also important. For this, write
Span $(v_1, \ldots, v_\ell)$ = set of all linear combinations of $v_1, \ldots, v_\ell$. Suppose V is finite dimensional.
Then choose $v_1 \neq 0$, $v_1 \in V$.
If span $(v_1) = V$, $\{v_1\}$ is a basis.
If span$(v_1) \neq V$, choose $v_2 \notin$ span(v_1).
Then v_1, v_2 are linearly independent.
If span$(v_1, v_2) = V$, v_1, v_2 are a basis.
If span$(v_1, v_2) \neq V$, choose $v_3 \notin$ span(v_1, v_2).
Then v_1, v_2, v_3 are linearly independent.
This process continued must stop eventually, since V finite dimensional, say generated by n vectors, implies that no more than n such independent $v_1, \ldots, v_\ell$ can be obtained. Thus at some stage, $v_1, \ldots, v_\ell$ are independent & span$(v_1, \ldots, v_\ell) = V$.

Note that this process, suitably varied, also shows that (assuming V finite dimensional) any independent set $v_1, \ldots, v_\ell$ can

be "extended" to be a basis, that is, $\exists\, v_{\ell+1} \cdots, v_k$ $k =$ dimension of V, such that

$$v_1, \ldots, v_\ell, v_{\ell+1}, \ldots, v_k \quad \text{is a basis.}$$

This is often important. In particular, it shows that if W is a subspace of V, then $\exists$ a basis of V of the form

$$v_1, \ldots, v_\ell, v_{\ell+1} \cdots v_k,$$

$\ell =$ dimension W such that $v_1, \ldots, v_\ell$ is a basis of W.

[Note: W is necessarily finite dimensional here because otherwise the expanding search for a basis of W, as we did on the previous page, would not terminate, W would contain linearly independent sets with arbitrarily many vectors. But since no set in V of more than $k = \dim V$ vectors can be independent, the expanding search for W stops with no more than k vectors: $\dim W \leq \dim V$ with equality only if $W = V$.]

The Rank & Nullity Theorem: If $T: V \to W$ is a linear transformation and if V is finite dimensional, then $\ker T$ and $\operatorname{im} T$ are finite dimensional and

$$\dim(\ker T) + \dim(\operatorname{im} T) = \dim V.$$

[Here as before $\ker T = \{v \in V : T(v) = \vec{0}_W\}$

and $\operatorname{im} T = \{Tv \in W : v \in V\}$, these being subspaces of V and W respectively].

Note that W need not be finite-dimensional here.

The proof will use the following two straightforward items, which are left as exercises

1. A linear transformation is injective (one to one) if and only if its kernel $= \{\vec{0}\}$
2. If T is an injective, surjective linear transformation from a finite-dimensional vector space V_1 to a vector space V_2, then V_2 is finite-dimensional and $\dim V_1 = \dim V_2$.

[Further exercises: 3. If $T: V_1 \to V_2$ is surjective and V_1 is finite dimensional, then V_2 is finite dimensional and $\dim V_2 \le \dim V_1$.
4. If $T: V_1 \to V_2$ is injective and V_2 is finite-dimensional, then V_1 is finite dimensional and $\dim V_1 \le \dim V_2$]

Proof of Rank & Nullity Theorem:
$\ker T \subset V$ so $\ker T$ is finite dimensional.
Choose a basis $v_1, \ldots, v_\ell$ $\quad \ell = \dim \ker T$
of $\ker T$ and extend this to be a basis
$v_1, \ldots, v_\ell, v_{\ell+1}, \ldots v_k$ $\quad k = \dim V$ of V,

as before. Set $\mathcal{U} = \operatorname{span}(v_{\ell+1}, \ldots, v_k)$.
Let $S : \mathcal{U} \to \operatorname{Im} T$ be the restriction of T to $\mathcal{U}$. Then
S is injective and surjective.
Surjective: $T\left(\sum_{i=1}^{k} \alpha_i v_i\right) = T\left(\sum_{i=\ell+1}^{k} \alpha_i v_i\right)$

since $T\left(\sum_{i=1}^{\ell} \alpha_i v_i\right) = 0$ and $T\left(\sum_{i=\ell+1}^{k} \alpha_i v_i\right) = S\left(\sum_{i=\ell+1}^{k} \alpha_i v_i\right)$.

Injective: $\ker S = \mathcal{U} \cap \ker T$
But $u \in \mathcal{U}$ has form $\sum_{i=\ell+1}^{k} \alpha_i v_i$ while

$u \in \ker T$ has form $\sum_{i=1}^{\ell} \alpha_i v_i$. By the

uniqueness of representation of a vector relative to a basis (which follows immediately from independence), these two can be equal only if all α's $= 0$, i.e. $u = \vec{0}$.
By the exercise above applied to S,

dimension $\operatorname{Im} T$ = dimension of $\mathcal{U} = k - \ell$.
$\square$

Important example: If $T: V \to V$, V finite dimensional, then T is injective if and only if T is surjective.

Proof: inj $\Rightarrow \ker T = 0 \Rightarrow \dim \operatorname{Im} T = \dim V - 0$

$\Rightarrow \operatorname{Im} T = V \Rightarrow$ surj. Also surj $\Rightarrow \operatorname{Im} T = V \Longrightarrow \dim \ker T = \dim V - \dim \operatorname{Im} T$ $= \dim V - \dim V = 0 \Rightarrow \ker T = \{\vec{0}\} \Rightarrow$ inj. $\square$

A linear transformation $T: V \to W$ that is injective and surjective is invertible as a function, and its inverse is necessarily a linear transformation (check this for yourself). Such "invertible linear transformations" or "isomorphisms" will be important as we go along.

Exercise: The invertible linear transformations from V to V form a group. (V need not be finite-dimensional for this).

If V is not finite dimensional, a linear transformation from V to V can be injective without being surjective and surjective without being injective.

Ex: Let V = space of eventually 0 sequences

Maps: $(x_1, x_2, x_3 \ldots) \longrightarrow (0, x_1, x_2, x_3 \ldots)$ inj not surj

$(x_1, x_2, x_3 \ldots) \longrightarrow (x_2, x_3, x_4 \ldots)$ surj, not inj