NOTES FOR 29 AUG (THURSDAY)

1. Recap

(1) Proved the change of basis formula.
(2) Proved that elementary row operations do not change the column rank of a column echelon matrix. (I incorrectly concluded that hence this statement holds for all matrices and hence the row and column ranks are equal. Sorry.) Actually, the proof for a general matrix is much simpler : Suppose $R_i \to aR_i + bR_j$ (where $a \neq 0$). Let $A_1, \ldots, A_k$ form a basis for the column space. Then, if the new columns $A'_1, A'_2, \ldots$ are linearly dependent, $\sum_\beta A_{\alpha\beta} c_\beta = 0 \; \forall \; \alpha \neq i$ and $\sum_\beta (aA_{i\beta} + bA_{j\beta}) c_\beta = 0$ and hence $\sum_\beta aA_{i\beta} c_\beta = 0 \Rightarrow \sum_\beta A_{i\beta} c_\beta = 0$. Since the original columns were linearly independent, $c_i = 0$ - a contradiction. Therefore, the dimension of the column space is $\geq$ the original one. Since row operations are invertible, equality holds. Note that whilst bringing a column echelon matrix to a row echelon one, one does not change the column echelon property. Hence, one can do both operations to bring it to a standard form indicated in the class.
(3) Showed how to prove that vectors in $\mathbb{F}^n$ are linearly independent etc, algorithmically.

2. Linear transformations

Recall the definition of a linear map $T : V \to W$. Here is a proposition (whose proof is easy).

Proposition 2.1. *A map $T : V \to W$ is linear iff $T(cv + w) = cT(v) + T(w) \; \forall \; c \in \mathbb{F}, \; v, w \in V$.*

Proof. One direction is trivial. As for the other, $T(\vec{0} + \vec{0}) = \vec{0} + T(\vec{0})$ and hence $T(\vec{0}) = \vec{0}$. Now $T(c_1 v + c_2 w) = c_1 T(v) + T(c_2 w + \vec{0}) = c_1 T(v) + c_2 T(w)$. □

Also inductively, $T(\sum_i c_i v_i) = \sum_i c_i T(v_i)$. Here are examples and non-examples.

(1) The zero map and the identity map are linear transformations from V to V.
(2) Let V be the space of abstract polynomials over $\mathbb{F}$. Then $T : V \to V$ given by $Tp(x) = p'(x)$ is a linear map. Likewise, for the space of polynomial functions. (More generally, we can define abstract polynomials $\mathbb{F}[x_1, \ldots, x_n]$ of n variables as abstract polynomials with coefficients in a commutative ring (a commutative ring is almost like a field except that non-zero elements are not necessarily invertible. Clearly polynomials in any number of variables over a field are commutative rings. We can construct polynomial rings with coefficients in any commutative ring (after all we never used the ability to divide in their construction).) $\mathbb{F}[x_1, x_2, \ldots, x_{n-1}]$, and define the partial derivative linear maps.) Here is a useful little lemma.

Lemma 2.2. *An abstract polynomial $p(x_1, \ldots, x_n)$ over a field of characteristic zero is determined by its polynomial function. Moreover, if the field is $\mathbb{R}$ (or $\mathbb{C}$) and the polynomial function vanishes in a neighbourhood of a, it vanishes everywhere (and hence the corresponding abstract polynomial is also zero).*

Proof. The claim is that $p(x_1, \ldots, x_n) = \sum_\alpha D^{(\alpha)} p(a) \frac{(x-a)^\alpha}{\alpha_1! \alpha_2! \ldots}$ as an abstract polynomial. Indeed, firstly, translations are isomorphisms from the vector space of abstract polynomials to itself (exercise). Secondly, let $p(x_1, \ldots, x_n) = \sum_\alpha a_\alpha (x - a)^\alpha$, then $D^{(\alpha)} p(a) = a_\alpha \alpha_1! \alpha_2! \ldots$. (Where did

we use the fact that the field has characteristic zero?)
If the field is $\mathbb{R}$ (or $\mathbb{C}$), then the derivatives at a are determined by the function in an arbitrarily small neighbourhood of a (what does a complex derivative mean? That is a longer story, but one can write everything in terms of the real and imaginary parts and apply the same argument). $\square$

(3) The map $T(f(x)) = \int_0^1 f(x)dx$ is a linear map between continuous functions on $[0,1]$ and $\mathbb{R}$.
(4) Let A be a fixed $m \times n$ matrix. Then $T(X) = AX$ is a linear map from $\mathbb{F}^n$ to $\mathbb{F}^m$. Likewise, so is $T(X) = PXQ$ where P is an $m \times m$ matrix and Q is an $n \times n$ matrix.
(5) The map $T(x) = x^2$ from $\mathbb{R}$ to $\mathbb{R}$ is not linear.
(6) The map $T(x) = ax + b$ from $\mathbb{R}$ to $\mathbb{R}$ is not linear (despite the common terminology) if $b \neq 0$. The correct term is "affine".

We have the following fundamental but easy result.

Proposition 2.3. *Let V be a finite-dimensional vector space over a field and $\mathcal{B} = \{e_1, \ldots, e_n\}$ be an ordered basis. Let W be another vector space over the same field and let $w_1, \ldots, w_n$ be vectors in W. There is a unique linear transformation $T : V \to W$ such that $Te_j = w_j$.*

Indeed, defining $Te_j = w_j$, then we can define $T(v) = T(\sum v_i e_i) = \sum v_i w_i$. Clearly T is a linear map. If U is another such linear map, then $U(v) = U(\sum v_i e_i) = v_i w_i = T(v)$.

Here is a fundamental example that we will deal with in more detail later : Let $T : \mathbb{F}^n \to \mathbb{F}^m$ be a linear transformation. It is determined uniquely by $\beta_i = T(e_i)$. Indeed, if $v = (v_1, v_2, \ldots) = \sum_j v_j e_j$, then $T(v) = \sum_j v_j \beta_j = \sum_i v_j \beta_{ij} e_i$ and hence $T(\vec{v}) = B\vec{v}$ where the $m \times n$ matrix B is $B_{ij} = (\beta_j)_i$.

Note that if $T : V \to W$ is linear, then $Ran(T) \subset W$ is a subspace (easy to see). Its dimension is called the rank of T. Likewise, it is easy to see that all $v \in V$ such that $T(v) = 0$ is a subspace of V. It is called the null space or sometimes, the kernel of T. Its dimension is called the nullity of T. Here is an important result.

Theorem 2.4. *$rank(T) + nullity(T) = dim(V) = n$.*

Proof. Let $e_1, \ldots, e_r$ be a basis for the kernel $null(T)$. Extend this set to a basis $e_1, \ldots, e_r, f_1, \ldots, f_{n-r}$. Then,

(1) $T(f_i)$ span the range : Indeed, $T(v) = \sum_i v_i T(e_i) + \sum_j w_j T(f_j) = \sum_j w_j T(f_j)$.
(2) $T(f_i)$ are linearly independent : If $\sum_i c_i T(f_i) = 0$, then $T(\sum_i c_i f_i) = 0$ and hence $\sum_i c_i f_i = \sum_k d_k e_k$. Thus, $c_i = 0 \ \forall\ i$.

Hence $T(f_i)$ form a basis for the range and we are done. $\square$

Here is another proof of the row rank being the column rank.

Proof. Consider the linear map $T : \mathbb{F}^n \to \mathbb{F}^m$ given by $T(v) = Av$. The range of T is the column space of T. Hence, $null(T) = n - c$ where c is the column rank. Bringing A to its row echelon form we see that we can solve for r variables (where r is the row rank) in terms of the $n - r$ free variables. Hence, $ker(T) = n - r$. Thus, $r = c$. $\square$

Actually, the proof above shows that $r = c$ is equivalent to the nullity-rank theorem.

This is a good point to introduce a few abstract constructions of vector spaces from existing spaces.
Def : Let V_i be an arbitrary collection of (not necessarily finite-dimensional) vector spaces.

(1) The direct product $\times_i V_i$: As a set it is simply the (possibly infinite) Cartesian product. The vector space structure is as follows : $\vec{0}$ is the element all of whose components are 0. Addition

and scalar multiplication is done component-wise. Additive inverses are also component wise. (More formally, the Cartesian product is the set of functions $f : I \to \cup_i V_i$ such that $f(i) \in V_i$. So for instance, the zero vector is the function $\vec{0}(i) = \vec{0}_i \in V_i$.)

(2) The direct sum $\oplus_i V_i \subset \times_i V_i$: It is a subspace of the direct product given by elements such that all but finitely many components are zero. (It is clearly closed under addition and scalar multiplication.)

While the direct product may look natural, actually the direct sum is better behaved. Indeed, if e_{ij} is a basis for V_i, then $f_{ij}(k) = e_{ij}\delta_{ki}$ is a basis for $\oplus_i V_i$ but a similar construction does not work for the direct product in general. (Obviously, because the direct sum is a proper subspace for infinite products.) Finite direct products and sums are the same (even for infinite-dimensional vector spaces). The internal direct sum of two subspaces is isomorphic to their direct sum (abstractly).