A note on the existence of determinants.

Since I ran out of time when finishing this proof in class today, I'm writing it up. (I think I might also have hurriedly switched an i and a k in class.)

Let F be a field, and recall, we defined the determinant as follows:

(1) For $n = 1$, we define $D : (F^1)^1 \to F$ by $D(\alpha) = \alpha$.

(2) For $n > 1$, we assume we have already defined a determinant $D : (F^{n-1})^{n-1} \to F$ satisfying the axioms. We now define $D : (F^n)^n \to F$ as follows: Let $v_1, \dots, v_n \in F^n$, and let $a_{i,j} \in F$ be such that $v_j = \sum_i a_{i,j} e_i$. Let $v_1', \dots, v_n'$ be the vectors $v_1, \dots, v_n$ (respectively) with each of their first coordinates deleted. Explicitly, $v_i' = \begin{pmatrix} a_{2,i} \\ \vdots \\ a_{n,i} \end{pmatrix}$.

Then we define $D(v_1, \dots v_n)$ by

$$D(v_1, \dots, v_n) = \sum_{j=1}^{n} (-1)^{1+j} a_{1,j} D(v_1', \dots, \overbrace{v_{j-1}', v_{j+1}'}^{v_j' \text{ omitted}}, \dots v_n')$$

Note that on the right hand side, each of the vectors have $n-1$ coordinates. Also in the j^{th} term, D only has $n-1$ arguments, since we omit v_j. Thus the right hand side only involves using $D : (F^{n-1})^{n-1} \to F$, which we have (inductively) assumed exists.

Remark 1. The definition given above is called the expansion about the first row. If you look at the proof, you'll see that we can just as well expand about any other row. Indeed, if you pick any i, and define $P_i \in \mathcal{L}(F^n, F^{n-1})$ to be the transformation that 'deletes' the i^{th} coordinate of v. Now the determinant expansion about the i^{th} row will read

$$D(v_1, \dots, v_n) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} D(P_i v_1, \dots, \overbrace{P_i v_{j-1}, P_i v_{j+1}}^{P v_j \text{ omitted}}, \dots P_i v_n).$$

Now the proof showing that D defined by expanding about the first row is a determinant is *almost identical* to the proof showing D defined by expanding about the i^{th} row is a determinant. The only place there is a minor difference in the proofs is checking $D(e_1, \dots, e_n) = 1$, but that's easy enough. Thus, by uniqueness of determinants, expanding the determinant about *any* row gives the same answer. We'll eventually see that you can do this about any column too.

Now to prove that the function D above is a determinant, we need to check that it maps the identity to 1, scales linearly if any argument is scaled, and remains unchanged if one argument is replaced with the sum of itself and another one. We will instead check that D satisfies the following properties:

(1) $D(e_1, \dots, e_n) = 1$

(2) $\forall \alpha \in F$, $D(\alpha v_1, v_2, \dots, v_n) = \alpha D(v_1, \dots, v_n)$.

(3) $D(v_1 + v_2, v_2, \dots, v_n) = D(v_1, \dots v_n)$.

(4) *(Swaps)* $\forall i < k$, we have

$$D(v_1, \dots, \overbrace{v_{i-1}, v_k, v_{i+1}}^{v_k \text{ in } i^{\text{th}} \text{ position}}, \dots, \overbrace{v_{k-1}, v_i, v_{k+1}}^{v_i \text{ in } k^{\text{th}} \text{ position}}, \dots v_n) = -D(v_1, \dots, v_n).$$

Note the difference between the above list, and the standard axioms: The standard axioms allow you to scale any argument, not only the first, as assumed above. The standard axioms also allow you to add any argument to any other argument, not only the second one to the first one, as assumed above. However, the last property above states that swapping arguments introduces a minus sign, and so, with the assumptions above we can immediately prove that scaling any argument (not just the first), scales the entire determinant by the same factor. Similarly we can use the above properties to prove that adding any argument to any other argument does not change the determinant.

Now the first three properties were proved in class. Here's a proof of the last one.

Proof that swapping arguments introduces a negative sign. Assume as above $v_j = \sum_i a_{i,j} e_i$. Now, by definition of D we have

$$(1)\quad D(v_1, \dots, \overbrace{v_{i-1}, v_k, v_{i+1}}^{v_k \text{ in } i^{\text{th}} \text{ position}}, \dots, \overbrace{v_{k-1}, v_i, v_{k+1}}^{v_i \text{ in } k^{\text{th}} \text{ position}}, \dots v_n)$$

$$= \sum_{j \neq i,k} (-1)^{1+j} a_{1,j} D(v'_1, \dots, \overbrace{v'_{i-1}, v'_k, v'_{i+1}}^{v'_k \text{ in } i^{\text{th}} \text{ position}}, \dots, \overbrace{v'_{k-1}, v'_i, v'_{k+1}}^{v'_i \text{ in } k^{\text{th}} \text{ position}}, \dots, \overbrace{v'_{j-1}, v'_{j+1}}^{v'_j \text{ omitted}}, \dots, v'_n)$$

$$+ (-1)^{1+i} a_{1,k} D(v'_1, \dots, \underbrace{v'_{i-1}, v'_{i+1}}_{v'_k \text{ in } i^{\text{th}} \text{ position omitted}}, \dots, \overbrace{v'_{k-1}, v'_i, v'_{k+1}}^{v'_i \text{ in } k^{\text{th}} \text{ position}}, \dots, v'_n)$$

$$+ (-1)^{1+k} a_{1,i} D(v'_1, \dots, \overbrace{v'_{i-1}, v'_k, v'_{i+1}}^{v'_k \text{ in } i^{\text{th}} \text{ position}}, \dots, \underbrace{v'_{k-1}, v'_{k+1}}_{v'_i \text{ in } k^{\text{th}} \text{ position omitted}}, \dots, v'_n)$$

Now note that by the inductive hypothesis, swapping arguments of D *on the right hand side* introduces a negative sign. Thus we have

$$\sum_{j \neq i,k} (-1)^{1+j} a_{1,j} D(v'_1, \dots, \overbrace{v'_{i-1}, v'_k, v'_{i+1}}^{v'_k \text{ in } i^{\text{th}} \text{ position}}, \dots, \overbrace{v'_{k-1}, v'_i, v'_{k+1}}^{v'_i \text{ in } k^{\text{th}} \text{ position}}, \dots, \overbrace{v'_{j-1}, v'_{j+1}}^{v'_j \text{ omitted}}, \dots, v'_n)$$

$$= - \sum_{j \neq i,k} (-1)^{1+j} a_{1,j} D(v'_1, \dots, \overbrace{v'_{i-1}, v'_i, v'_{i+1}}^{v'_i \text{ in } i^{\text{th}} \text{ position}}, \dots, \overbrace{v'_{k-1}, v'_k, v'_{k+1}}^{v'_k \text{ in } k^{\text{th}} \text{ position}}, \dots, \overbrace{v'_{j-1}, v'_{j+1}}^{v'_j \text{ omitted}}, \dots, v'_n)$$

$$= - \sum_{j \neq i,k} (-1)^{1+j} a_{1,j} D(v'_1, \dots, \overbrace{v'_{j-1}, v'_{j+1}}^{v'_j \text{ omitted}}, \dots, v'_n)$$

For the term

$$D(v'_1, \dots, \underbrace{v'_{i-1}, v'_{i+1}}_{v'_k \text{ in } i^{\text{th}} \text{ position omitted}}, \dots, \overbrace{v'_{k-1}, v'_i, v'_{k+1}}^{v'_i \text{ in } k^{\text{th}} \text{ position}}, \dots, v'_n)$$

we note that v'_k never occurs as an argument, and that v'_i is 'out of place'. Swapping v'_i with it's previous argument $(k-1) - (i+1) - 1$ times will put it 'back in place'. Thus

$$D(v'_1, \dots, \underbrace{v'_{i-1}, v'_{i+1}}_{v'_k \text{ in } i^{\text{th}} \text{ position omitted}}, \dots, \overbrace{v'_{k-1}, v'_i, v'_{k+1}}^{v'_i \text{ in } k^{\text{th}} \text{ position}}, \dots, v'_n)$$

$$= (-1)^{k-i-3} D(v'_1, \dots, \overbrace{v'_{i-1}, v'_i, v'_{i+1}}^{v'_i \text{ in } i^{\text{th}} \text{ position}}, \dots, \underbrace{v'_{k-1}, v'_{k+1}}_{v'_k \text{ in } k^{\text{th}} \text{ position omitted}}, \dots, v'_n)$$

$$= -(-1)^{k-i} D(v'_1, \dots, \underbrace{v'_{k-1}, v'_{k+1}}_{v'_k \text{ in } k^{\text{th}} \text{ position omitted}}, \dots, v'_n)$$

Similarly for the last term on the right of equation (1), swapping v'_k with it's next argument $(k-1)-(i+1)-1$ times will put it in the k^{th} position. Thus

$$D(v'_1, \ldots, \overbrace{v'_{i-1}, v'_k, v'_{i+1}}^{v'_k \text{ in } i^{\text{th}} \text{ position}}, \ldots, \underbrace{v'_{k-1}, v'_{k+1}}_{v'_i \text{ in } k^{\text{th}} \text{ position omitted}}, \ldots, v'_n) = -(-1)^{k-i} D(v'_1, \ldots, \underbrace{v'_{i-1}, v'_{i+1}}_{v'_i \text{ in } i^{\text{th}} \text{ position omitted}}, \ldots, v'_n)$$

Finally collecting all these simplified expressions and substituting back in (1) gives

$$\begin{aligned} D(v_1, \ldots, &\overbrace{v_{i-1}, v_k, v_{i+1}}^{v_k \text{ in } i^{\text{th}} \text{ position}}, \ldots, \overbrace{v_{k-1}, v_i, v_{k+1}}^{v_i \text{ in } k^{\text{th}} \text{ position}}, \ldots v_n) \\ &= -\sum_{j \neq i,k} (-1)^{1+j} a_{1,j} D(v'_1, \ldots, \overbrace{v'_{j-1}, v'_{j+1}}^{v'_j \text{ omitted}}, \ldots, v'_n) \\ &\quad - (-1)^{1+i} a_{1,k} (-1)^{k-i} D(v'_1, \ldots, \underbrace{v'_{k-1}, v'_{k+1}}_{v'_k \text{ in } k^{\text{th}} \text{ position omitted}}, \ldots, v'_n) \\ &\quad - (-1)^{1+k} a_{1,i} (-1)^{k-i} D(v'_1, \ldots, \underbrace{v'_{i-1}, v'_{i+1}}_{v'_i \text{ in } i^{\text{th}} \text{ position omitted}}, \ldots, v'_n) \end{aligned}$$

Since $(-1)^{1+2k-i} = (-1)^{1+i}$, the right hand side of the above reduces exactly to $-D(v_1, \ldots, v_n)$, finishing the proof. $\square$