Solution for HW5, part D
Geoff Anderson
granders@fas.harvard.edu

Problem 7

Recall that we have an isomporhism of vector spaces $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$.

(a) Consider the determinant map $det : M_n(\mathbb{R}) \to \mathbb{R}$, and find $\nabla(det)(A)$, expressed in terms of $A = [a_{ij}]$.

(b) Consider the function $f : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ given by $f(A) = A^2$. Show that $Jf_A(H) = AH + HA$.

Solution

If we have a function $f : \mathbb{R}^{n^2} \to \mathbb{R}$, ∇f (evaluated at a point) is an element of $\mathbb{R}^{n^2}$. So we can think of $\nabla det(A)$ as an element of $M_n(\mathbb{R})$.

Let's compute the ijth component of $\nabla det(A)$. Using the cofactor expansion along the ith row, we can write

$$det(A) = \sum_{k=1}^{n} (-1)^{(i+k)} a_{ik} \cdot det(A_{ik}).$$

(Where A_{ij} denote the $(n-1)$ by $(n-1)$ submatrix obtained by deleting the ith row and jth column from A.) There is only one term in this sum containing a_{ij}, and $det(A_{ij})$ is a polynomial in the terms a_{lm} where $l \neq i$ and $m \neq j$. Hence if we take the ijth partial derivative, we get $D_{ij}det(A) = (-1)^{(i+j)}det(A_{ij})$.

Hence

$$\nabla det(A) = [(-1)^{(i+j)}det(A_{ij})].$$

If A is not invertible, this is the best we can do. If A is invertible on the other hand, there is a nicer way to express this, if we use the following idenity:

$$A^{-1} = (1/det(A)) \cdot [(-1)^{(i+j)}det(A_{ji})].$$

Then we have

$$\nabla det(A) = [D_{ij}det(A)] = [(-1)^{(i+j)}det(A_{ij})] = det(A) \cdot (A^{-1})^t.$$

You can check this identity as an exercise, or look it up... it (worded slightly differently) along with proof can be found in Schneider and Barker, theorem 4.5.10, page 200. It follows more or less directly from proposition 4.5.7 on page 198. In their notation, $\Gamma_{ij} = (-1)^{(i+j)}det(A_{ij})$ is called the ijth *cofactor* (it is the coefficient of a_{ij} in the expression for $det(A)$).

(b) Recall that that the Jacobian is the *unique* linear map satisfying

$$\lim_{|H| \to 0} \frac{|f(A+H) - f(A) - Jf_A(H)|}{|H|} = 0$$

It suffices therefore to show that the map $L : M_n \to M_n$ defined by $L(H) = AH + HA$ is linear and satisfies ‡:

$$\lim_{|H| \to 0} \frac{|f(A+H) - f(A) - L(H)|}{|H|} = 0.$$

Linearity of L follows from distributivity of matrix multiplication. Let's check ‡.
$f(A+H) - f(A) = AH + HA + H^2$, so $f(A+H) - f(A) - L(H) = H^2$. Hence we need to show that $\lim_{|H|\to 0} |H^2|/|H| = 0$. Luckily for us, the inequality $|H^2| \leq |H|^2$ holds, so the limit is in fact 0. ($|H^2|$ is not the same as $|H|^2$ in general!)

This inequality can be derived from the following lemma. Lemma: let $A, B \in M_n(\mathbb{R})$. Then $|A \cdot B| \leq |A| \cdot |B|$. Proof: in what follows, let A_i be the ith *row* vector of A and B_j the jth *column* vector of B. Also keep in mind that $|A|^2 = \sum_i |A_i|^2$ and $|B|^2 = \sum_i |B_i|^2$.

$|A \cdot B|^2 = \sum_{i,j} < A_i, B_j >^2$ ($<, >$ the standard euclidean inner product on $\mathbb{R}^n$). The Schwartz inequality tells us $< A_i, B_j >^2 \leq |A_i|^2 |B_j|^2$. Furthermore, $|B_j|^2 \leq |B|$ for all j. Hence,

$$\sum_{i,j} < A_i, B_j >^2 \leq \sum_{i,j} |A_i|^2 |B_j|^2 \leq |B|^2 \cdot \sum_i |A_i|^2 = |A|^2 |B|^2$$

and we get the desired inequality.