Differential Geometry—MTG 6256—Fall 2012
Problem Set 1: Fun with Matrices

Below, $M_n(\mathbf{R})$ denotes the vector space of $n \times n$ real matrices, and $\mathrm{GL}(n, \mathbf{R})$ the subset of invertible matrices.

The problems below are meant to be done in the given order; in many cases the results of earlier problems are applicable to later problems. When asked to find the derivative of a map, express your answer by writing down a formula that gives all directional derivatives.

1. Let $U \in \mathbf{R}$ be open, W a finite-dimensional vector space, $f : U \to W$ differentiable at $p \in U$. Show that $(D_p f)(1) = f'(p)$.

2. Let V, W_1, W_2 be finite-dimensional vector spaces, $U \subset V$ open, $p \in U$, and $g_i : U \to W_i$ differentiable at p for $i = 1, 2$. Define $f : U \to W_1 \oplus W_2$ by $f(q) = (g_1(q), g_2(q))$. Show that f is differentiable at p and compute $Df|_p(v)$ for arbitrary $v \in V$.

Note: one general approach to a problem of the form "show that a function F is differentiable at point q, and compute the derivative $DF|_q$" is to compute all the directional derivatives $(D_q F)(v)$. If this expression is not linear in v, then F is not differentiable at q (and you were instructed to show something that was false). If "$(D_q F)(v)$" *is* linear in v, then the linear transformation T defined by $T(v) = (D_q F)(v)$ is the *only candidate* for $DF|_q$. You can then try to show that F is differentiable at q either by plugging this T into the definition of "differentiable at q" and showing that the relevant limit is zero, or by showing that, for all fixed v, the map $\tilde{q} \mapsto (D_{\tilde{q}} F)(v)$ is continuous in $\tilde{q}$ (in which case, automatically, F is not merely differentiable at q, but continuously differentiable at q).

3. Define $\mu : M_n(\mathbf{R}) \oplus M_n(\mathbf{R}) \to M_n(\mathbf{R})$ by $\mu(A, B) = AB$ (matrix multiplication). Show that μ is differentiable, and compute its derivative.

4. Let $g, h : M_n(\mathbf{R}) \to M_n(\mathbf{R})$ be differentiable and define $f(A) = g(A)h(A)$. Note that $f = \mu \circ j$, where μ is the map in problem 3 and $j : M_n(\mathbf{R}) \to M_n(\mathbf{R}) \oplus M_n(\mathbf{R})$ is defined by $j(A) = (g(A), h(A))$. Prove that f is differentiable, and (using directional derivatives) express the derivative of f in terms of the derivatives of g and h.

If your answer is correct, then in the case $n = 1$, you should find with the aid of problem 1 that you've recovered the "product rule" from Calculus 1. Thus, the Calculus-1 product rule is a corollary of the (multivariable) Chain Rule.

5. Define $f : M_n(\mathbf{R}) \to M_n(\mathbf{R})$ by $f(A) = A^t A$, where A^t is the transpose of A. Show that f is differentiable and find its derivative.

6. Let $m \geq 1$ be an integer and let $f(A) = A^m$ for $A \in M_n(\mathbf{R})$. Show that f is differentiable and find its derivative.

7. We saw in class that $\mathrm{GL}(n, \mathbf{R})$ is an open subset of $M_n(\mathbf{R})$. Show that it is also dense. Hint for a quick proof: characteristic polynomial. (Look up the term if you've forgotten what it means, or never learned it.) Of course, there are many other proofs as well.

8. Define $\iota : \mathrm{GL}(n, \mathbf{R}) \to M_n(\mathbf{R})$ by $\iota(A) = A^{-1}$. In this exercise, you will show that ι is differentiable without using the power-series approach we used in class, and without needing to show ahead of time that ι is continuous. (In class, I used the continuity of ι to deduce that directional derivatives were continuous, and said that continuity is not hard to show. That fact is true; we just don't need it in the approach below. Then we can deduce continuity of ι from the general fact that differentiability implies continuity.) The idea is the following:

- For a fixed A, we find a linear transformation T that is *the only possible candidate* T for the derivative of ι at A.
- We plug this T into the quotient whose limit we take in the definition of "derivative", and show directly that the limit of the quotient is zero.

The same method can be used in many other examples. Usually the candidate T is found by computing directional derivatives, but in the case of ι there is a "cheaper" approach, which is part (a) below.

(a) Using the result of problem 4, show that <u>if</u> ι is differentiable, then $(D_A\iota)(B) = -A^{-1}BA^{-1}$ for all $A \in M_n(\mathbf{R}), B \in \mathrm{GL}(n, \mathbf{R})$. Note that for fixed A, this is linear in B.

(b) Show that if $A \in M_n(\mathbf{R})$, then A is invertible if and only if A is *bounded below*, i.e. iff there exists $c > 0$ such that $\|Av\| \geq c\|v\|$ for all $v \in \mathbf{R}^n$.

(c) Using part (b) and the triangle inequality, show that for all $A \in \mathrm{GL}(n, \mathbf{R})$, there exists $\delta > 0$ such that if $\|B\| < \delta$ then $A + B$ is bounded below, hence is invertible. (This gives another proof that $\mathrm{GL}(n, \mathbf{R})$ is open in $M_n(\mathbf{R})$, of course.) Here and below, the norm used on $M_n(\mathbf{R})$ is the operator norm.

(d) Fix $A \in \mathrm{GL}(n, \mathbf{R})$ and define T to be the linear transformation found above in part (a), the map $B \mapsto -A^{-1}BA^{-1}$. Using just algebraic manipulation and the submultiplicativity of the operator norm, show that $\|\iota(A + B) - \iota(A) - T(B)\| \leq \|(A + B)^{-1}\| \; \|A^{-1}\|^2 \; \|B\|^2$.

(e) Show that if A and c are as in part (b), then $\|A^{-1}\| \leq \frac{1}{c}$.
Note: Since $1 = \|I\| = \|AA^{-1}\| \leq \|A\| \; \|A^{-1}\|$, we have the simple *lower* bound $\|A^{-1}\| \geq \|A\|^{-1}$. But there is no general *upper* bound on $\|A^{-1}\|$ in terms of $\|A\|$.

(f) Use part (e) to show that your work in part (c) actually gives you, for fixed A, a uniform upper bound on $\|(A + B)^{-1}\|$ if $\|B\|$ is sufficiently small.

(g) Now show that if $A \in \mathrm{GL}(n, \mathbf{R})$ then $\lim_{B \to 0} \frac{\|\iota(A+B)-\iota(A)-T(B)\|}{\|B\|} = 0$, hence that ι is differentiable at A.

9. Extend the result of problem 6 to negative integral exponents. (For $A \in \mathrm{GL}(n, \mathbf{R})$ and $m \geq 1$, A^{-m} is defined to be $(A^{-1})^m$.)

10. The determinant function $\det : M_n(\mathbf{R}) \to \mathbf{R}$ is a polynomial in n^2 variables, so it is certainly C^1 (in fact C^∞). There are several ways to compute its derivative. The steps below constitute a method that involves little computation but a bit of thought.

(a) Let $I \in M_n(\mathbf{R})$ be the identity and let $B \in M_n(\mathbf{R})$. Compute $D_I(\det)(B)$, and express the answer as a simple invariant of the matrix B.

(b) Let $A \in \mathrm{GL}(n, \mathbf{R}), B \in M_n(\mathbf{R})$. Compute $D_A(\det)(B)$. (Hint: use (a).) Re-express your result as a formula for the function $\log|\det|$.

(c) Use the density statement in problem 7 to extend the formula for $D_A(\det)$ from $A \in \mathrm{GL}(n, \mathbf{R})$ to $A \in M_n(\mathbf{R})$. The answer can be rewritten in terms of the "cofactor" matrix $\mathrm{cof}(A)$ that arises in computing the inverse of a matrix. (Recall that if A is invertible, then $A^{-1} = \frac{1}{\det(A)}\mathrm{cof}(A)$, or else the transpose of this, depending on your definition of $\mathrm{cof}(A)$.)