MATH 23A 2ND MIDTERM SOLUTIONS

1. (George) a) $A \subset \mathbf{R}^n$ is closed if and only if for all sequences $\{x_i\} \subset A$ such that $\{x_i\}$ converges to x in $\mathbf{R}^n$, $\{x_i\}$ converges in A (that is, $x \in A$).

   b) Suppose A is closed according to the above definition. Consider $y \in \mathbf{R}^n$ such that there is no ϵ such that $B_\epsilon(y) \subset \mathbf{R}^n \setminus A$. Thus for any i there is an $x_i \in A$ with $|x_i - y| < 1/i$. It is clear that $\{x_i\}$ converges to y in $\mathbf{R}^n$ so by the definition above y is in A. Therefore for every z in $\mathbf{R}^n \setminus A$ there is an ϵ such that $B_\epsilon(z) \subset \mathbf{R}^n \setminus A$ so $\mathbf{R}^n \setminus A$ is open.

   Suppose $\mathbf{R}^n \setminus A$ is open and $\{x_i\}$ converges to x in $\mathbf{R}^n$ and has terms in A. Given $\epsilon > 0$ there is N such that $x_i \in B_\epsilon(x)$ for all $i > N$. In particular, there is at least one element of A in $B_\epsilon(x)$ so no neighborhood of x lies in $\mathbf{R}^n \setminus A$ and since $\mathbf{R}^n \setminus A$ is open that means that $x \in A$.

   c) Suppose f is continuous and A is a closed set in $\mathbf{R}$. Consider a sequence $\{x_i\}$ that converges to x in $\mathbf{R}^n$ such that each x_i is in $f^{-1}(A)$. Since f is continuous we know that $\lim(f(x_i)) = f(\lim(x_i))$ so $f(x) = \lim(f(x_i))$. It is clear that the sequence $\{f(x_i)\}$ is in A and since A is closed so is $\lim(f(x_i)) = f(x)$. Therefore $x \in f^{-1}(A)$ so $f^{-1}(A)$ is closed.

   Now suppose $f^{-1}(A)$ is open for every open set A. Given $x \in \mathbf{R}^n$ and $\epsilon > 0$ consider $f^{-1}(B_\epsilon(f(x))) \subset \mathbf{R}^n$. Since $(B_\epsilon(f(x))$ is open in R $f^{-1}(B_\epsilon(f(x)))$ is open in $\mathbf{R}^n$ so there is a $\delta > 0$ such that
$$B_\delta(x) \subset f^{-1}(B_\epsilon(f(x)))$$
   This means that $|x - y| < \delta$ (i.e. $y \in B_\delta(x)$) implies that $|f(x) - f(y)| < \epsilon$ (i.e. $f(y) \in B_\epsilon(f(x))$). This is exactly the desired result.

2. (Matt) Define $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ by $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{x+y}{2} \\ \frac{x-y}{2} \\ xy \end{pmatrix}$. Define $g\colon \mathbb{R} \to \mathbb{R}^2$ by $g(t) = \begin{pmatrix} e^t \\ e^{-t} \end{pmatrix}$. Find $\mathrm{D}(f \circ g)(t)$.

   1) using the Chain Rule

   f and g have continuous partial derivatives, so their derivatives are just their Jacobian matrices:
$$\mathrm{D}f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \mathrm{D}_1 f_1 & \mathrm{D}_2 f_1 \\ \mathrm{D}_1 f_2 & \mathrm{D}_2 f_2 \\ \mathrm{D}_1 f_3 & \mathrm{D}_2 f_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \\ y & x \end{pmatrix}$$
$$\mathrm{D}g(t) = \begin{pmatrix} \mathrm{D}_1 g \\ \mathrm{D}_2 g \end{pmatrix} = \begin{pmatrix} e^t \\ -e^{-t} \end{pmatrix}$$
   So we have $[\mathrm{D}(f \circ g)(t)] = [\mathrm{D}f(g(t))] \circ [\mathrm{D}g(t)] = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \\ e^{-t} & e^t \end{pmatrix} \begin{pmatrix} e^t \\ -e^{-t} \end{pmatrix} =$
$$\mathrm{D}(f \circ g)(t) = \begin{pmatrix} \frac{e^t - e^{-t}}{2} \\ \frac{e^t + e^{-t}}{2} \\ 0 \end{pmatrix}.$$
   2) by directly differentiating $(f \circ g)$. First we compute $(f \circ g)$:

$$(f \circ g)(t) = \begin{pmatrix} \frac{e^t+e^{-t}}{2} \\ \frac{e^t-e^{-t}}{2} \\ 1 \end{pmatrix}.$$

Then we determine the partial derivatives of $(f \circ g)$, find that they are all continuous, and use them to write the Jacobian matrix:

$$\mathrm{D}(f \circ g)(t) = \begin{pmatrix} \frac{e^t-e^{-t}}{2} \\ \frac{e^t+e^{-t}}{2} \\ 0 \end{pmatrix}.$$

We see that the Chain Rule yields the same answer as does direct differentiation.

3. (David) Let $F : M_n \to M_n$ be the map given by $F(A) = AA^T$.

a)Prove that F is continuous.

To prove that F is continuous at $A \in M_n$ we have to show that $|F(A + H) - F(A)|$ is small if $|H|$ is small. By the definition

$$F(A + H) - F(A) = (A + H)(A^T + H^T) - AA^T = AH^T + A^T H + HH^T$$

Therefore

$$|F(A+H) - F(A)| = |AH^T + A^T H + HH^T| \leq |AH^T| + |A^T H| + |HH^T| \leq |A||H^T| + |A^T||H| + |H||H^T| \leq 2|A||H| + |H|^2 \tag{1}$$

So $|F(A + H) - F(A)|$ is small if $|H|$ is small.

b)Prove that F is differentiable and find $D(F)(A) : M_n \to M_n$ for $A \in M_n$.

To prove that F is differentiable at $A \in M_n$ we have to show a linear map $C : M_n \to M_n$ such that for small $H \in M_n$ we have

$$|F(A + H) - F(A) - C(H)|/|H| \to 0 for H \to 0$$

. In this case $D(F)(A) = C$. Define a linear map $C : M_n \to M_n$ by $C(H) = AH^T + A^T H$. Then $F(A+H) - F(A) - C(H) = HH^T$ and therefore $|F(A+H) - F(A) - C(H)|/|H| = |HH^T|/|H| \leq |H|$. So F is differentiable at A and the linear map $D(F)(A) : M_n \to M_n$ is given by $D(F)(A)(H) = AH^T + A^T H$

4. (Nils) Recall from problem set 5/solution set 5 that any choice of basis on $V = \mathbf{R}^n$ gives an isomorphism of bilinear forms on $\mathbf{R}^n$ with $n \times n$ matrices, i.e., $\mathcal{B}(\mathbf{R}^n) \cong \mathrm{Mat}(n, n)$. Since a bilinear form is antisymmetric iff the matrix in one/any basis is antisymmetric, this means that we also get $\mathcal{B}_{an}(\mathbf{R}^n) \cong \mathrm{Mat}_{an}(n, n)$, that is, the space of antisymmetric bilinear forms on $\mathbf{R}^n$ is isomorphic (given a choice of basis) with the space of antisymmetric $n \times n$ matrices. To compute the dimension of this latter space, note that every antisymmetric $n \times n$ matrix is of the form:

$$\begin{pmatrix} 0 & a_1 & a_2 & \cdots\cdots & \\ -a_1 & 0 & a_n & \cdots\cdots & \\ \cdots & \ddots & \ddots & \ddots & \cdots \\ \cdots\cdots & & \ddots & 0 & a_N \\ \cdots\cdots\cdots & & & -a_N & 0 \end{pmatrix}, \tag{2}$$

where[1] $N = \sum_{i=1}^{n-1} i = \frac{(n-1)n}{2} = \binom{n}{2}$, so $\dim \mathcal{B}_{an}(\mathbf{R}^n) = \binom{n}{2}$. In particular,

(a) For $n = 1$, $N = 0$.

[1]Recall that the binomial coefficient $\binom{n}{k} = \frac{n!}{(n-k)!k!}$.

(b) For $n = 2$, $N = 1$.
(c) For $n = 3$, $N = 3$.
(d) For $n = 10$, $N = 45$.

Note in fact that one can "pull" the direct sum decomposition $\mathcal{B}(\mathbf{R}^n) = \mathcal{B}_s(\mathbf{R}^n) \oplus \mathcal{B}_{an}(\mathbf{R}^n)$ over to matrices, so $\mathrm{Mat}(n,n) = \mathrm{Mat}_s(n,n) \oplus \mathrm{Mat}_{an}(n,n)$. This gives $\dim \mathcal{B}_s(\mathbf{R}^n) = n^2 - \binom{n}{2} = \binom{n+1}{2}$.

5. (Li-chung)

(a) We claim that the partial derivatives of f are continous. Let $F_1, F_2, \ldots, F_n$ be the component functions of F, i.e. $F(u) = (F_1(u), F_2(u), \ldots, F_n(u))$. Then all the partial derivatives of $F_1, \ldots, F_n$ are continuous because F is continuously differentiable. Let $v = (v_1, v_2, \ldots, v_n)$. Then $f(u)$ is simply

$$(F_1(u) - v_1)^2 + (F_2(u) - v_2)^2 + \cdots + (F_n(u) - v_n)^2.$$

So for each $j = 1, \ldots, n$,

$$D_j f(u) = 2(F_1(u) - v_1) \cdot D_j F_1(u) + \cdots + 2(F_n(u) - v_n) \cdot D_j F_n(u).$$

This is a continuous function because $D_j F_i$ is continous for each i (as a function of u), F_i is continuous for each i (since F being differentiable implies each F_i is continuous), and $D_j F$ is just a polynomial expression of continuous functions. (Use theorem 1.5.27 repeatedly, and also note that constant functions are continuous.)

Because f has continuous partial derivatives everywhere, f is differentiable everywhere.

(b) We will use the formula of $D_j f$ derived above.

$$\begin{aligned}
& Df(u) = 0 \\
\iff & D_1 f(u) = D_2 f(u) = \cdots = D_n f(u) = 0 \\
\iff & \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} D_1 f(u) \\ D_2 f(u) \\ \vdots \\ D_n f(u) \end{bmatrix} = \begin{bmatrix} 2(F_1(u) - v_1) \cdot D_1 F_1(u) + \cdots + 2(F_n(u) - v_n) \cdot D_1 F_n(u) \\ 2(F_1(u) - v_1) \cdot D_2 F_1(u) + \cdots + 2(F_n(u) - v_n) \cdot D_2 F_n(u) \\ \vdots \\ 2(F_1(u) - v_1) \cdot D_n F_1(u) + \cdots + 2(F_n(u) - v_n) \cdot D_n F_n(u) \end{bmatrix} \\
= & \begin{bmatrix} D_1F_1(u) & D_1F_2(u) & \ldots & D_1F_n(u) \\ D_2F_1(u) & D_2F_2(u) & \ldots & D_2F_n(u) \\ \vdots & \vdots & & \vdots \\ D_nF_1(u) & D_nF_2(u) & \ldots & D_nF_n(u) \end{bmatrix} \begin{bmatrix} 2(F_1(u) - v_1) \\ 2(F_2(u) - v_2) \\ \vdots \\ 2(F_n(u) - v_n) \end{bmatrix} \\
= & DF(u)^T \begin{bmatrix} 2(F_1(u) - v_1) \\ 2(F_2(u) - v_2) \\ \vdots \\ 2(F_n(u) - v_n) \end{bmatrix}.
\end{aligned}$$

Now $DF(u)$ is invertible, so $DF(u)^T$ is too; specifically, $(DF(u)^{-1})^T$ is the inverse because $DF(u)^T (DF(u)^{-1})^T = \left(DF(u)^{-1} DF(u)\right)^T = I^T = I$ and $(DF(u)^{-1})^T DF(u)^T = \left(DF(u) DF(u)^{-1}\right)^T = I^T = I$. Hence $DF(u)T$ as a linear transformation $\mathbf{R}^n \to \mathbf{R}^n$ is an isomorphism.

This implies that $\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = DF(u)^T \begin{bmatrix} 2(F_1(u) - v_1) \\ 2(F_2(u) - v_2) \\ \vdots \\ 2(F_n(u) - v_n) \end{bmatrix}$ occurs if and only if $\begin{bmatrix} 2(F_1(u) - v_1) \\ 2(F_2(u) - v_2) \\ \vdots \\ 2(F_n(u) - v_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, which occurs if and only if $F_i(u) = v_i$ for all i. This is the same as saying $F(u) = v$.