PRELIM II: SOLUTIONS.

1. Let $A \in Mat_{n\times n}(\mathbb{R})$.

   a. TRUE: If λ is an eigenvalue of A, then λ^k is an eigenvalue of A^k (for $k \geq 1$ a positive integer).

   Proof: By induction on k.
   Base case: $k=1$. We know λ^1 is an eigenval of A^1.
   Induction step: Suppose λ^k is an eigenvalue of A^k.
   Notationally, let $\vec{v}$ be a non-zero vector satisfying $A\cdot\vec{v} = \lambda\cdot\vec{v}$ and $A^k\cdot\vec{v} = \lambda^k\cdot\vec{v}$.
   Then
   $$\begin{aligned} A^{k+1}\vec{v} &= A\cdot(A^k\vec{v}) \\ &= A\cdot(\lambda^k\vec{v}) && \text{by induction;} \\ &= \lambda^k(A\cdot\vec{v}) && \text{because } A \text{ is linear;} \\ &= \lambda^k(\lambda\vec{v}) && \text{because } A\vec{v}=\lambda\vec{v}; \\ &= \lambda^{k+1}\vec{v}, && \text{as desired.} \end{aligned}$$ $\square$

   b. FALSE: If $\lambda \geq 0$ is an eigenvalue of A^2, $\sqrt{\lambda}$ need not be an eigenvalue of A.

   For example, if $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, then
   $$A\vec{v} = -\vec{v}$$
   for every non-zero $\vec{v}$, so its only eigenvalue is -1.

   On the other hand, $A^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and so
   $$A^2\vec{v} = \vec{v}$$
   for every non-zero $\vec{v}$. So the only eigenvalue of A^2 is 1. But $-1 \neq +\sqrt{1}$, so the statement is FALSE.

2. a. $GL_2(\mathbb{R})$ is an open subset of $\mathbb{R}^4$. See the proof of Corollary 1.5.39 in our text for full details.

   b. $GL_2(\mathbb{R})$ is NOT closed in $\mathbb{R}^4$, because its complement is NOT open. For example, $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^4 - GL_2(\mathbb{R})$, but any ε-ball about $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ contains $\begin{bmatrix} \varepsilon/2 & 0 \\ 0 & \varepsilon/2 \end{bmatrix}$, which is invertible, hence in $GL_2(\mathbb{R})$ and NOT in $\mathbb{R}^4 - GL_2(\mathbb{R})$. Thus, $GL_2(\mathbb{R})$ is not closed.

PRELIM II SOLUTIONS

2c. $$f: GL_2(\mathbb{R}) \to GL_2(\mathbb{R})$$
$$A \longmapsto A^{-1}$$

in $\mathbb{R}^4$ coordinates is
$$f\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} d \\ -b \\ -c \\ a \end{pmatrix}.$$

Each coordinate function is thus a rational function on $\mathbb{R}^4$, and as we know the denominator to be non-zero on $GL_2(\mathbb{R})$, this is a continuous function. Thus f itself is continuous. $\square$

3a. $\boxed{\text{TRUE}}$ Proof by contradiction. Let $a \neq b$, and suppose $f(b) = f(a)$. Then by the Mean Value theorem, there is a value c between a & b so that
$$\frac{f(b) - f(a)}{b-a} = f'(c).$$

But if $b \neq a$, then this says $f'(c) = 0$, contradicting our assumption. Hence f is one-to-one. $\square$

b. $\boxed{\text{FALSE}}$ Consider the function $f: \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$. Then $[Df\begin{pmatrix} x \\ y \end{pmatrix}] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, but f is <u>not</u> 1-1 since $f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = f\begin{pmatrix} 1 \\ 5 \end{pmatrix}$. $\square$

c. $\boxed{\text{FALSE}}$ Consider the function $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}$. Then $\left[Df\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix}$

and so $\det[Df] = e^x \cos^2 y + e^x \sin^2 y = e^x \neq 0$.
Thus, $[Df]$ is ~~not~~ invertible, but
$$f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = f\begin{pmatrix} 0 \\ 2\pi \end{pmatrix},$$ so f is NOT one-to-one $\square$

4.a. We may use Prop. 2.8.9 to compute a global Lipschitz ratio.

$$D_1F_1 = \cos(x-y) \qquad D_2F_1 = -\cos(x-y) + 2y$$
$$D_1F_2 = -\sin(x+y) - 1 \qquad D_2F_2 = -\sin(x+y)$$

$$D_1D_1F_1 = -\sin(x-y) \qquad D_1D_2F_1 = -\sin(x-y)$$
$$D_2D_1F_1 = -\sin(x-y) \qquad D_2D_2F_1 = +\sin(x-y) + 2$$

$$D_1D_1F_2 = -\cos(x+y) \qquad D_1D_2F_2 = -\cos(x+y)$$
$$D_2D_1F_2 = -\cos(x+y) \qquad D_2D_2F_2 = -\cos(x+y).$$

Each of these satisfies
$$|D_iD_jF_k(\vec{x})| \le 1$$
except $D_2D_2F_1$, which satisfies
$$|D_2D_2F_1(\vec{x})| \le 3.$$
Thus, by prop. 2.8.9,
$$|DF(\vec{u}) - DF(\vec{v})| \le \underbrace{\sqrt{1^2+3^2+1^2+1^2+1^2+1^2+1^2+1^2}}_{=\sqrt{16}\,=\,4}\,|u-v|$$

Thus, F has a global Lipschitz ratio of $\boxed{4}$.

b. Consider $G = F\begin{pmatrix}x\\y\end{pmatrix} - \begin{pmatrix}0.5\\0\end{pmatrix}$. Then $DG = DF$, and we want to apply Newton's method to G.

$$\vec{a}_1 = \vec{a}_0 - [DG(a_0)]^{-1}\,\vec{G}(\vec{a}_0)$$
$$= \begin{pmatrix}0\\0\end{pmatrix} - \begin{bmatrix}1 & -1\\ -1 & 0\end{bmatrix}^{-1}\cdot\begin{bmatrix}-0.5\\1\end{bmatrix}$$
$$= \begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}-0.5\\1\end{bmatrix} = \boxed{\begin{bmatrix}1\\1/2\end{bmatrix}}$$

c. We want to check if we may apply Kantorovich's Thm. Thus, we want to check the inequality (2.8.51),
$$|G(\vec{a}_0)|\cdot|[DG(a_0)]^{-1}|^2 \le \frac{1}{2}.$$

We have

$$|G(\vec{a}_0)| \left|[DG(\vec{a}_0)]^{-1}\right|^2 = \left|\begin{bmatrix} -0.5 \\ 1 \end{bmatrix}\right| \cdot \left|\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right|^2$$

$$= \left(\sqrt{\tfrac{5}{4}}\right) \cdot 3 = \frac{3\sqrt{5}}{2} \geq \frac{3}{2} > \frac{1}{2}.$$

Thus, Kantorovich's Theorem does not apply; so we cannot say if Newton's method converges.

SOME OF YOU programmed this into a computer to do a few more steps of Newton's Method, and found that after 3 ~~or~~ 4 steps, the criterion is satisfied, Kantorovich's Thm does then apply, and Newton's method DOES in fact converge. GOOD WORK!

5. First, we show directional derivatives exist at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

$$D_{\vec{v}} f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lim_{t \to 0} \frac{f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + t\vec{v}\right) - f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right)}{t}.$$

Now, $f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \cdot h(\vec{x}) = 0$ for any $\vec{x} \in S$, so $f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0$. Thus, the limit becomes

$$= \lim_{t \to 0} \frac{f(t \cdot \vec{v})}{t}$$

$$= \lim_{t \to 0} \frac{f\left(|t \cdot \vec{v}| \frac{t \cdot \vec{v}}{|t \cdot \vec{v}|}\right)}{t}$$

$$= \lim_{t \to 0} \frac{t \cdot |\vec{v}|\, h\left(\frac{|t| \cdot \vec{v}}{|t \cdot \vec{v}|}\right)}{t}$$

But now we notice $\frac{|t| \cdot (\vec{v})}{|t \cdot \vec{v}|} = \frac{\vec{v}}{|\vec{v}|}$, so this is

$$= \lim_{t \to 0} |\vec{v}|\, h\left(\frac{\vec{v}}{|\vec{v}|}\right)$$

$$= |\vec{v}|\, h\left(\frac{\vec{v}}{|\vec{v}|}\right) = f(\vec{v}).$$

Thus, $D_{\vec{v}} f\begin{pmatrix}0\\0\end{pmatrix} = f(\vec{v})$ exists for all $\vec{v} \neq \vec{0}$.

NOW we must show f is differentiable everywhere $\Leftrightarrow h\begin{pmatrix}x\\y\end{pmatrix} = ax+by$

($\Rightarrow$) If f is differentiable, it's differentiable at the origin. Then $\left[Df\begin{pmatrix}0\\0\end{pmatrix}\right]$ is its derivative, and that must be linear $\left[Df\begin{pmatrix}0\\0\end{pmatrix}\right] = [a \quad b]$.

But for $\vec{v} \in S$, we know

$$D_{\vec{v}} f\begin{pmatrix}0\\0\end{pmatrix} = f(\vec{v}) - h(\vec{v})$$

and also,

$$D_{\vec{v}} f\begin{pmatrix}0\\0\end{pmatrix} = \left[Df\begin{pmatrix}0\\0\end{pmatrix}\right] \cdot \vec{v}$$

SO, $h\begin{pmatrix}x\\y\end{pmatrix} = D_{\begin{pmatrix}x\\y\end{pmatrix}}\begin{pmatrix}0\\0\end{pmatrix} = [a \quad b]\begin{bmatrix}x\\y\end{bmatrix} = ax+by$, as desired.

($\Leftarrow$) We know that if $h\begin{pmatrix}x\\y\end{pmatrix} = ax+by$, then $f\begin{pmatrix}x\\y\end{pmatrix} = ax+by$ is thus linear, and linear functions are differentiable. $\square$

6. First, we know linear functions are C^1, and sums of C^1 functions are C^1, so $f = g+T$ is C^1.

Next, we note $\left|g\begin{pmatrix}0\\0\end{pmatrix}\right| \leq M|\vec{0}|^2 = 0$, so $g\begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$.

Moreover,

$$\lim_{|\vec{h}|\to 0} \frac{g(\vec{0}+\vec{h}) - g(\vec{0})}{|\vec{h}|} = \lim_{|\vec{h}|\to 0} \frac{g(\vec{h})}{|\vec{h}|}.$$

But $\lim_{\vec{h}\to\vec{0}} \left|\frac{g(\vec{h})}{|\vec{h}|}\right| \leq \lim_{\vec{h}\to\vec{0}} \frac{|g(\vec{h})|}{|\vec{h}|} \leq \lim_{\vec{h}\to\vec{0}} M \frac{|\vec{h}|^2}{|\vec{h}|} = \lim_{\vec{h}\to\vec{0}} M|\vec{h}| = 0$

Absolute convergence implies convergence, so

$$\left[Dg\begin{pmatrix}0\\0\end{pmatrix}\right] = \underline{0}.$$

is the zero linear transformation.

Thus, $[Df] = [Dg] + T = T$. But $T: \mathbb{R}^n \to \mathbb{R}^n$ is surjective, hence invertible, so $[Df(\vec{0})]$ is invertible. By the Inverse Function Theorem, then, f is invertible in a neighborhood of $\vec{0}$. $\square$

7.a. Note that $A^T A$ has (i,j)-entry $\vec{v}_i \cdot \vec{v}_j$, where $\vec{v}_k$ denotes the k^{th} column of A. THUS,

$$A \in O(n) \iff \text{Columns of } A \text{ are an orthonormal basis of } \mathbb{R}^n$$
$$\iff \vec{v}_i \cdot \vec{v}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$
$$\iff A^T A = I. \quad \square$$

b. Suppose $A, B \in O(n)$. Then

(i) $$(AB)^T(AB) = B^T A^T A B = B^T I B = B^T B = I \Rightarrow AB \in O(n).$$

(ii) $$A^T A = I \Rightarrow A^T = A^{-1} \Rightarrow AA^T = I \Rightarrow (A^T)^T A^T = I \Rightarrow A^T = A^{-1} \in O(n). \quad \square$$

c. For any matrix A,

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

so $A^T A \in S(n)$.

d. Similarly to Example ~~1.7.18~~ 1.7.18, for $F(A) = A^T A$,

$$[DF(A)] \cdot H = H^T A + A^T H.$$

To show this is onto when A is invertible, choose $B \in S(n)$, and let $H = (A^{-1})^T \frac{B}{2}$. Then

$$H^T A + A^T H = \left((A^{-1})^T \frac{B}{2}\right)^T A + A^T \left((A^{-1})^T \frac{B}{2}\right)$$
$$= \frac{B^T}{2} A^{-1} A + A^T \cdot (A^T)^{-1} \frac{B}{2}.$$
$$= \frac{B}{2} + \frac{B}{2} = B.$$

Thus, $[DF(A)]$ surjects onto $S(n)$. $\square$

PRELIM II SOLUTIONS

e. We want to apply Theorem 3.1.16.

$$F : \mathrm{Mat}_{n\times n}(\mathbb{R}) \longrightarrow S(n)$$
$$A \longmapsto A^T A$$

is certainly a C^1 mapping, as shown in (d). Moreover, $I \in S(n)$ is a point, so certainly a manifold. ~~[illegible]~~ Moreover,

$$O(n) = F^{-1}(I)$$

by (a), and (d) shows $[DF(A)]$ is onto for each $A \in F^{-1}(I)$. Thus, $O(n) = F^{-1}(I)$ is a manifold.

To determine the tangent space at I, we note

$$O(n) = \{G(A) = \vec{0} \mid G(A) = A^T A - I\}.$$

~~[illegible]~~ So

$$\begin{aligned} T_I O(n) &= \ker([DG(A)]) \\ &= \ker([DF(A)]) \\ &= \{H \mid H^T I + I^T H = 0\} \\ &= \{H \mid H^T + H = 0\} \\ &= \{H \mid H^T = -H\} \\ &= A(n). \quad \square \end{aligned}$$

f. We know if $AB = BA$, $e^A e^B = e^{A+B}$. If $A \in A(n)$, then $A^T = -A$. Moreover, $(A^m)^T = (A^T)^m$ (CHECK THIS!), so $(e^A)^T = e^{(A^T)}$. Thus, for $A \in A(n)$,

$$(e^A)(e^A)^T = e^A e^{A^T} = e^A e^{-A} = e^{A-A} = e^0 = I.$$

Thus, $e^A \in O(n)$. $\square$

g. For $A \in A(2)$, $A = \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}$. It is not hard to compute

$$e^A = \begin{pmatrix} \cos(a) & -\sin(a) \\ \sin(a) & \cos(a) \end{pmatrix}$$

is a <u>rotation</u> matrix. NOTE exp is <u>NOT</u> onto $O(2)$!