Problem Set #4

(<u>1.7.15</u>) (a) Define what it means for a mapping $F: Mat(m,n) \to Mat(k,\ell)$ to be differentiable at a point $A \in Mat(m,n)$.

Recall that we can identify an element of $Mat(m,n)$ with an element of $\mathbb{R}^{mn}$. Thus this is just the usual definition of differentiability:

F is differentiable if there exists a linear transformation $L_A : Mat(m,n) \to Mat(k,\ell)$ such that

$$\lim_{H \to 0} \frac{1}{|H|}\Big(F(A+H) - F(A) - L_A(H)\Big) = \vec{0} \in Mat(k,\ell)$$

The linear transformation might also be written as $[DF(A)]$ or $[JF(A)]$.

(b) Consider the function $F: Mat(n,m) \to Mat(n,m)$ given by $F(A) = AA^T$. Show that F is differentiable and compute the derivative $[DF(A)]$.

We will mirror the book's analysis from Example 1.7.18. Note that

$$\begin{aligned} F(A+H) - F(A) &= (A+H)(A+H)^T - AA^T \\ &= (A+H)(A^T + H^T) - AA^T \\ &= AA^T + AH^T + HA^T + HH^T - AA^T \\ &= AH^T + HA^T + HH^T \end{aligned}$$

Then we guess (and will check momentarily) that $[DF(A)]$ will be the linear part of this (linear in H), so guess

$$[DF(A)]H = AH^T + HA^T. \qquad (*)$$

First let's check that this $[DF(A)]$ is indeed a linear transformation:

$$\begin{aligned} \text{(i)}\ [DF(A)](H_1+H_2) &= A(H_1+H_2)^T + (H_1+H_2)A^T \\ &= AH_1^T + AH_2^T + H_1A^T + H_2A^T \\ &= [DF(A)]H_1 + [DF(A)]H_2 \end{aligned}$$

$$\begin{aligned} \text{(ii)}\ [DF(A)](\lambda H) &= A(\lambda H)^T + (\lambda H)A^T \\ &= \lambda(AH^T + HA^T) \\ &= \lambda [DF(A)]H \end{aligned}$$

Thus $[DF(A)]$ is indeed a linear transformation.

Now let's check that this $[DF(A)]$ which we guessed actually satisfies the definition of differentiability given in part (a).

(on next page) $\longrightarrow$

(1.7.15 (b) continued)

By absolute convergence, it is enough to show that

$$\lim_{H \to 0} \frac{1}{|H|} \left| \left(F(A+H) - F(A) - [DF(A)]H \right) \right| = 0$$

With our choice $[DF(A)]H = AH^T + HA^T$, observe that

$$\lim_{H \to 0} \frac{1}{|H|} \left| F(A+H) - F(A) - [DF(A)]H \right|$$

$$= \lim_{H \to 0} \frac{1}{|H|} \left| AH^T + HA^T + HH^T - AH^T - HA^T \right|$$

$$= \lim_{H \to 0} \frac{1}{|H|} \left| HH^T \right|$$

$$\leq \lim_{H \to 0} \frac{|H||H^T|}{|H|} \qquad \text{by Schwarz's Inequality}$$

$$= \lim_{H \to 0} |H^T|$$

$$= \lim_{H \to 0} |H| \qquad \text{since } |H^T| = |H|$$

$$= 0$$

Thus F is differentiable and $[DF(A)]H = AH^T + HA^T$.

(1.7.16) Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $A^2 = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$

(parts a,b,c)

(a) Write the formula for the function $S: \mathbb{R}^4 \to \mathbb{R}^4$ defined by

$$S\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{pmatrix}$$

We are effectively identifying $Mat(2,2)$ with $\mathbb{R}^4$ in the usual way and saying that S is the map which squares the matrix. Note that

$$A^2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2+bc & ab+bd \\ ca+dc & cb+d^2 \end{bmatrix}$$

So S is given by $S\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} a^2+bc \\ ab+bd \\ ca+dc \\ cb+d^2 \end{pmatrix}$

(b) Find the Jacobian matrix of S.

This is just the 4×4 matrix of partial derivatives:

$$[DS(A)] = \begin{bmatrix} 2a & c & b & 0 \\ b & a+d & 0 & b \\ c & 0 & a+d & c \\ 0 & c & b & 2d \end{bmatrix}$$

(c) Check that your answer agrees with example 1.7.18.

We will need to create a generic matrix $H = \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix} \longleftrightarrow \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix}$

and show that $[DS(A)]H = AH + HA$ (from the example).

$$AH+HA = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix} + \begin{bmatrix} h_1 & h_2 \\ h_3 & h_4 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ah_1+bh_3 & ah_2+bh_4 \\ ch_1+dh_3 & ch_2+dh_4 \end{bmatrix}$$
$$+ \begin{bmatrix} ah_1+ch_2 & bh_1+dh_2 \\ ah_3+ch_4 & bh_3+dh_4 \end{bmatrix} = \longrightarrow$$

(continued)

(1.7.16 (c) continued)

$$= \begin{bmatrix} 2ah_1 + bh_3 + ch_2 & ah_2 + b(h_1 + h_4) + dh_2 \\ ah_3 + c(h_1 + h_4) + dh_3 & bh_3 + ch_2 + 2dh_4 \end{bmatrix}$$

$$= \begin{bmatrix} 2ah_1 + bh_3 + ch_2 \\ ah_2 + b(h_1 + h_4) + dh_2 \\ ah_3 + c(h_1 + h_4) + dh_3 \\ bh_3 + ch_2 + 2dh_4 \end{bmatrix}$$ (written as a vector in $\mathbb{R}^4$)

Meanwhile, $[DS(A)]H = \begin{bmatrix} 2a & c & b & 0 \\ b & a+d & 0 & b \\ c & 0 & a+d & c \\ 0 & c & b & 2d \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix}$

$$= \begin{bmatrix} 2ah_1 + bh_3 + ch_2 \\ ah_2 + b(h_1 + h_4) + dh_2 \\ ah_3 + c(h_1 + h_4) + dh_3 \\ bh_3 + ch_2 + 2dh_4 \end{bmatrix}$$

Thus the two are indeed equal, verifying the example 1.7.18.

(1.7.18) In example 1.7.18, prove that the derivative $AH+HA$ is the same as the Jacobian matrix computed with partial derivatives.

This is precisely parts (b) and (c) of exercise 1.7.16 above, so there is no need to rewrite it.

(1.7.19) Is the mapping $f: \mathbb{R}^n \to \mathbb{R}^n$ given by $f(\bar{x}) = |\bar{x}|\bar{x}$ differentiable at the origin? If so, what is the derivative?

We say that f is differentiable at $\bar{a}$ if

$$\lim_{\bar{h} \to \bar{0}} \frac{1}{|\bar{h}|} \left| f(\bar{a} + \bar{h}) - f(\bar{a}) - [Df(\bar{a})]\bar{h} \right| = \bar{0}$$

for some linear transformation $[Df(\bar{a})]$, which we call the derivative.

Filling in $\bar{a} = \bar{0}$, this becomes

$$\lim_{\bar{h} \to \bar{0}} \frac{1}{|\bar{h}|} \left| f(\bar{h}) - f(\bar{0}) - [Df(\bar{a})]\bar{h} \right|$$

$$= \lim_{\bar{h} \to 0} \frac{1}{|\bar{h}|} \left| |\bar{h}|\bar{h} - Df(\bar{a})\bar{h} \right|$$

Note that
$$\lim_{\bar{h} \to 0} \frac{1}{|\bar{h}|} \left| |\bar{h}|\bar{h} \right| = \lim_{\bar{h} \to 0} \frac{|\bar{h}||\bar{h}|}{|\bar{h}|}$$
$$= \lim_{\bar{h} \to 0} |\bar{h}|$$
$$= 0$$

Thus, if we take $[Df(\bar{a})]$ to be the zero transformation given by
$$[Df(\bar{a})]\bar{h} = \bar{0}$$
then the limit is 0.

Hence f is differentiable and the derivative is the $n \times n$ matrix of all zero entries.