Chapter 6 Section 3

Lehner White

December 4, 2015

6.12

(i)

$$
\begin{aligned}
Df(x) &= DAx \\
&= D\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= D\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}
\end{aligned}
$$

Because we know that the matrix represntation of the linear map $Df(x)$ is the jacobian it follows that:

$$
\begin{aligned}
D\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} &= \begin{bmatrix} D_1f_1(x) & D_2f_1(x) & \cdots & D_nf_1(x) \\ D_1f_2(x) & D_2f_2(x) & \cdots & D_nf_2(x) \\ \vdots & \vdots & \ddots & \vdots \\ D_1f_m(x) & D_2f_m(x) & \cdots & D_nf_m(x) \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = A
\end{aligned}
$$

(ii)

$$\begin{aligned}
Df(x) &= Dx^T A \\
&= D\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \\
&= D\begin{bmatrix} a_{11}x_1 + \cdots + a_{m1}x_n & a_{12}x_1 + \cdots + a_{m2}x_n & \cdots & a_{1n}x_1 + \cdots + a_{mn}x_n \end{bmatrix} \\
&= D\begin{bmatrix} f_1 & f_2 & \cdots & f_n & \end{bmatrix}
\end{aligned}$$

Similar to part **(i)** it follows that:

$$D\begin{bmatrix} f_1 & f_2 & \cdots & f_n & \end{bmatrix} = \begin{bmatrix} D_1 f_1(x) & D_2 f_1(x) & \cdots & D_n f_1(x) \\ D_1 f_2(x) & D_2 f_2(x) & \cdots & D_n f_2(x) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f_m(x) & D_2 f_m(x) & \cdots & D_n f_m(x) \end{bmatrix}^T = A^T$$

6.13

(i)

If $f(x) = u(x)^T v(x)$ then we want to show that:

$$\lim_{h\to 0} \frac{\|u(x+h)^T v(x+h) - u(x)v(x) - u(x)^T Dv(x)h - v(x)^T Du(x)h\|}{\|h\|} = 0$$

And from proposition 6.2.17 we know that $\mathbf{u}, \mathbf{v}$ are locally lipschitz giving us that:

$$\|u(x+h) - u(x)\| \leq L\|h\| \quad \|v(x+h) - v(x)\| \leq L\|h\|$$

And it follows that:

$$\|u(x+h) - u(x)^T - Du(x)\| \leq \frac{\epsilon\|h\|}{3(\|v(x)\|+1)}$$

$$\|v(x+h) - v(x)^T - Dv(x)\| \leq \frac{\epsilon\|h\|}{3(\|u(x)\|+1)}$$

If $\|h\| < \delta_x$ then both dervivatives exists for some ball.

For $\epsilon > 0$ if $\delta = \min\left\{\delta_x, \frac{\epsilon}{3L\|Dv(x)\|}\right\}$ then $\|h\| < \delta$ gives us that:

$$\begin{aligned}
&\|u(x+h)^T v(x+h) - u(x)v(x) - u(x)^T Dv(x)h - v(x)^T Du(x)h\| \\
&\quad \leq \|u(x+h)^T\|\|v(x+h) - v(x) - Dv(x)h\| + \|v(x)\|\|u(x+h) - u(x)^t - Du(x)h\| \\
&\qquad\quad + \|u(x+h)^T - u(x)^T\|\|Dv(x)h\|\|h\| \\
&\quad \leq \|u(x)^T + L\|\|v(x+h) - v(x) - Dv(x)h\| + \|v(x)\|\|u(x+h)^T - u(x)^T - Du(x)h\| \\
&\qquad\quad + \delta L\|Dv(x)\|\|h\| \\
&\quad < \epsilon\|h\|
\end{aligned}$$

(ii)
If $f(x) = x^T \quad g(x) = Ax$ then it follows from part **(i)** that:

$$D(fg) = x^t A + x^T A^T = x^T(A + A^T)$$

(iii)
If $f(x) = Bw$ then it follows from part **(i)** that:

$$\begin{aligned} Df(x) &= B(x)Dw(x) + w^T DB(x)^T \\ &= B(x)Dw(x) + \begin{bmatrix} w^T Db_1^T(x) \\ \vdots \\ w^T(x)Db_k^T(x) \end{bmatrix} \end{aligned}$$

6.14

Finding the standard inner product we have that:

$$\begin{aligned} \int \overline{DF(\gamma(t))^T}\gamma'(t)dt &= \int DF(\gamma(t))\gamma'(t)dt \\ &= \int DF(\gamma(t))D\gamma(t)\gamma'(t)dt \\ &= \int DC \cdot D\gamma(t)\gamma'(t)dt \\ &= \int 0 \cdot D\gamma(t)\gamma'(t)dt \\ &= 0 \end{aligned}$$

And since the inner product is 0 we know that they are orthogonal.

6.15

We know that:

$$D(g \circ f) = D(g(f(\mathbf{x})))D(f(\mathbf{x}))$$

$$D(g(x)) = \begin{bmatrix} y & x \\ 2x & 2y \end{bmatrix}$$

$$D(f(x)) = \begin{bmatrix} -1 & cos(y) \\ e^x & -1 \end{bmatrix}$$

We also know that Because $f(0,0) = (0,1)$ we know that:

$$Dg(f(0,0)) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$$

6.16

Taking the inner product here gives us:

$$\begin{aligned}\langle f, g\rangle &= xy(\sin(y) - x) + (e^x - y)(x^2 + y^2) \\ &= xy\sin(y) - x^2y + x^2e^x + y^2e^x - yx^2 - y^3\end{aligned}$$

And taking the derivative of this function will be the gradient giving:

$$D(\langle f, g\rangle) = \begin{bmatrix} y\sin(y) - 2xy + 2xe^x + x^2e^x + y^2e^x - 2xy \\ x\sin(y) + xy\cos(y) - x^2 + 2ye^x - x^2 - 3y^2 \end{bmatrix}$$

And evaluating at $(0, 0)$ gives us $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

6.17

If $f(\mathbf{x}) = \|A\mathbf{x} - \mathbf{b}\|_2^2$ using the chain rule here we have that:

$$Df(\mathbf{x}_0) = 2\|A\mathbf{x}_0 - \mathbf{b}\|_2 \cdot \nabla\|A\mathbf{x}_0 - \mathbf{b}\|_2 \cdot A$$