MATH 208: Homework #9

Jesse Farmer

08 March 2004

1. *In what direction is the largest value of the directional derivative assumed and what is that largest value?*

   From the following problem we know that $D_v f(x_0) = \nabla f(x_0) \cdot v$. Therefore, because v is a unit vector,

   $$\nabla f(x_0) \cdot v = \|v\|\|\nabla f(x_0)\| \cos \theta = \|\nabla f(x_0)\| \cos \theta$$

   which attains its maximum when $\theta = 0$, assuming we are only considering $\theta \in [0, 2\pi)$, i.e., $\nabla f(x_0)$ and v point in the same direction. Moreover the maximum is clearly $\|\nabla f(x_0)\|$.

2. *Prove that $D_v f(x_0) = \nabla f(x_0) \cdot v$.*

   In this case, we have $f : \mathbb{R}^n \to \mathbb{R}$ differentiable at x_0, since otherwise $\nabla f(x_0)$ does not make any sense with our definition. Recall that $\nabla f(x_0) = \big(D_1 f(x_0), \ldots, D_n f(x_0)\big) \in \mathbb{R}^n$. First we show that $D_v f(x_0) = Df(x_0)v$.

   Let $h = tv$ and consider

   $$0 = \lim_{t \to 0} \frac{|f(x_0 + tv) - f(x_0) - Df(tv)|}{|tv|} = \frac{1}{|v|} \left(\lim_{t \to 0} \frac{|f(x_0 + tv) - f(x_0) - tDf(x_0)|}{|t|} \right)$$

   which proves this statement.

   Then

   $$
   \begin{aligned}
   D_v f(x_0) &= Df(x_0)v \\
   &= Df(x_0)(v_1 e_1 + \cdots + v_n e_n) \\
   &= \sum_{j=1}^{n} v_j Df(x_0)(e_j) \\
   &= \sum_{j=1}^{n} v_j D_j f(x_0) \\
   &= \nabla f(x_0) \cdot v
   \end{aligned}
   $$

3. *Are any of $A_4, D_{12}, S_3 \times \mathbb{Z}_2$ isomorphic?*

We know that $D_6 \cong S_3$ and that $D_6 \times \mathbb{Z}_2 \cong D_{12}$, so $S_3 \times \mathbb{Z}_2 \cong D_{12}$. This problem is from one of our exams last quarter. Consider the following table,

A_4	Order	D_{12}	Order
(1)	1	e	1
(123)	3	r	6
(132)	3	r^2	3
(124)	3	r^3	2
(142)	3	r^4	3
(134)	3	r^5	6
(143)	3	f	2
(234)	3	fr	2
(243)	3	fr^2	2
$(12)(34)$	2	fr^3	2
$(13)(24)$	2	fr^4	2
$(14)(23)$	2	fr^5	2

Since A_4 has no elements of order greater than 3 and D_{12} has two it follows that they cannot be isomorphic. Therefore $A_4 \not\cong D_{12} \cong S_3 \times \mathbb{Z}_2$.

4. *Show that there exists a unique $a \in A^+, n \in N, k \in O(n, \mathbb{R})$ such that $g = kan$ for any $g \in GL_n(\mathbb{R})$.*

We will begin by showing that every upper-triangular matrix can be uniquely decomposed into a matrix from A^+ and N, and then show that any matrix in $GL_n(\mathbb{R})$ can be uniquely decomposed into an upper-triangular matrix and an orthogonal matrix.

Let

$$U = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

be arbitrary. Then for $A \in A^+$ let $\lambda_i = a_{ii}$ and for $B \in N$ let $b_{ij} = \frac{a_{ij}}{\lambda_i}$. That is, $\lambda_i b_{ij} = a_{ij}$.

We then get

$$
\begin{aligned}
AB &= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \begin{pmatrix} 1 & b_{12} & \cdots & b_{1n} \\ 0 & 1 & & b_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & \lambda_1 b_{12} & \cdots & \lambda_1 b_{1n} \\ 0 & \lambda_2 & & \lambda_2 b_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \\
&= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} \\
&= U
\end{aligned}
$$

Looking at the multiplication from above, we see that if there were some other decomposition, $A'B'$, then $\lambda_i = \lambda_i'$, so $\lambda_i b_{ij} = \lambda_i b_{ij}' \Rightarrow b_{ij} = b_{ij}'$, i.e., the decomposition is unique.

Now it is sufficient to show that any non-singular matrix can be uniquely decomposed into an upper-triangular and an orthogonal matrix.

Given an orthonormal basis $\{v_1, \ldots, v_m\}$ of some vector space V with dimension m over $\mathbb{R}^n$ we want to find a unique $w \in V$ such that for any $x \in \mathbb{R}^n$ we have $x - w \in V^\perp$, i.e., $x - w$ is orthogonal to every vector in V. Since $w \in V$ there exists $\alpha_1, \cdots, \alpha_m$ such that

$$
w = \alpha_1 v_1 + \cdots + \alpha_m v_m
$$

So for any $1 \le i \le m$ we need

$$
\begin{aligned}
v_i \cdot (x - w) &= v_i \cdot (x - \alpha_1 v_1 - \cdots - \alpha_i v_i - \cdots - \alpha_m v_m) \\
&= v_i \cdot x - \alpha_1 (v_i \cdot v_1) - \cdots \alpha_i (v_i \cdot v_i) - \cdots - \alpha_m (v_i \cdot v_m) \\
&= v_i \cdot x - \alpha_i = 0
\end{aligned}
$$

This holds if and only if $\alpha_i = v_i \cdot x$. Therefore the unique w for which we are looking for is

$$
w = (v_1 \cdot x) v_1 + \cdots + (v_m \cdot x) v_m
$$

This is precisely the orthogonal projection of x onto V, denoted $\operatorname{proj}_V x$. Since any basis of a vector space can be represented as an invertible matrix, it follows that if we convert an arbitrary basis into an orthonormal basis we will be able to decompose the invertible matrix. We claim that this decomposition is into an orthogonal matrix and an upper-triangular matrix.

Let V be a vector space over $\mathbb{R}^n$ with dimension m and V_i the vector space generated by the first i basis vectors. Let $\{v_1, \ldots, v_m\}$ be a basis for V and consider V_1. We define

$$
w_1 = \frac{v_1}{\|v_1\|}
$$

This is clearly an orthonormal basis for V_1. Inductively then, assume that we have an orthonormal basis $\{w_1, \ldots, w_{i-1}\}$ for V_{i-1}. Define

$$w_i = \frac{v_i - \text{proj}_{V_{i-1}} v_i}{\|v_i - \text{proj}_{V_{i-1}} v_i\|}$$

From the statements about the orthogonal projection above we see that w_i is indeed orthogonal to each v_i and orthonormal since $\|w_i\| = 1$.

We must now express each v_i as a linear combination of w_i. We get

$$v_1 = \|v_1\| w_1$$

and

$$\begin{aligned} v_i &= \text{proj}_{V_{i-1}} v_i + \|v_i - \text{proj}_{V_{i-1}}\| w_i \\ &= (w_1 \cdot v_i) w_1 + \cdots + (w_{i-1} \cdot v_i) w_{i-1} + \|v_i - \text{proj}_{V_{i-1}}\| w_i \end{aligned}$$

For $i = 2, \ldots, m$ we let $a_{11} = \|v_1\|$, $r_{ii} = \|v_i - \text{proj}_{V_{i-1}}\|$, and $r_{ji} = w_j \cdot v_i$ and $i > j$.

Finally, let $A \in GL_n(\mathbb{R})$. Since A is non-singular it has linearly independent columns, denoted $v_1, \ldots, v_m$. Using the above process we can find an orthogonal matrix and an upper-triangular matrix such that

$$\begin{pmatrix} v_1 & \cdots & v_m \end{pmatrix} = \begin{pmatrix} w_1 & \cdots & w_m \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ 0 & r_{22} & & r_{2m} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & r_{mm} \end{pmatrix}$$

Therefore any non-singular matrix can be decomposed into an orthogonal matrix and an upper-triangular matrix and all that remains to be shown is that this decomposition is unique.

Let $A \in GL_n(\mathbb{R})$ be arbitrary and assume there exist $Q_1, Q_2 \in O(n, \mathbb{R})$ and $R_1, R_2 \in U$, the set of upper-triangular matrices such that $A = Q_1 R_1 = Q_2 R_2$. Since A is non-singular, both of these matrices are and

$$Q_2^{-1} Q_1 = R_2 R_1^{-1}$$

We know that the product of upper-triangular matrices or orthogonal matrices is upper-triangular or orthogonal, respectively, and likewise with their inverses. So both of these matrices are upper-triangular and orthogonal. Moreover, we know that R_2 has a positive determinant by construction, and hence that R_1^{-1} has a positive determinant. Since these are upper-triangular matrices this means that their diagonal is positive.

Let $B = (b_{ij})$ be both upper-triangular and orthogonal, then we know $B^T = B^{-1}$, so

$$\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & b_{nn} \end{pmatrix} \begin{pmatrix} b_{nn} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ b_{1n} & \cdots & b_{11} \end{pmatrix} = I$$

But B^T must be an upper-triangular matrix since the inverse of an upper-triangular matrix is upper-triangular, and hence B^T must be a diagonal matrix. So B must have a positive diagonal and, from the orthogonality of B, it follows that $B = I$.

Therefore $R_2 R_1^{-1} = I \Rightarrow R_1 = R_2$, and $Q_2^{-1} Q_1 = I \Rightarrow Q_1 = Q_2$. Therefore the decomposition is unique. Since both this and the decomposition of upper-triangular matrices is unique, our statement is proved.

5. *Prove that for $f : \mathbb{R}^n \to \mathbb{R}^m$ differentiable $Df(x) = \big(D_j f_i(x)\big)$.*

   We know that $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable if and only if each of its component functions is. Assuming f is differentiable at some point $x_0 \in \mathbb{R}^n$ then we also know that

   $$Df(x_0)v = \big(Df_1(x_0)v, \dots Df_n(x_0)v\big)$$

   Using Problem 2 above we get

   $$
   \begin{aligned}
   Df(x_0)v &= \begin{pmatrix} Df_1(a)v \\ \vdots \\ Df_n(a)v \end{pmatrix} \\
   &= \begin{pmatrix} \sum_{j=1}^n D_j f_1(x_0)v_j \\ \vdots \\ \sum_{j=1}^n D_j f_n(x_0)v_j \end{pmatrix} \\
   &= \big(D_j f_i(x_0)\big) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}
   \end{aligned}
   $$

6. *Prove that if $U \subseteq \mathbb{R}^n$ is open and connected and $f : U \to \mathbb{R}^m$ is differentiable with $Df(x) = 0$ for $x \in U$, then f is a constant function.*

   We first establish this face for $f : \mathbb{R} \to \mathbb{R}$. Let f be differentiable on some interval and $f'(x) = 0$ for all x in that interval. Let $a \neq b$ be any two points in the interval. Then by the mean value theorem there exists some $c \in (a, b)$ such that

   $$0 = f'(c) = \frac{f(b) - f(a)}{b - a}$$

   which implies that $f(a) = f(b)$ for all a, b in the interval, i.e., f is constant on that interval.

   Now let $U \subseteq \mathbb{R}^n$ be open and connected and $f : U \to \mathbb{R}^m$ a differentiable map such that $Df(x) = 0$ for all $x \in U$. Since f is constant if and only if each of its component functions is, and $Df(x) = 0$ as a matrix if and only if each of its rows is 0, we may assume that f is real-valued. Suppose that $\nabla f(x) = 0$ for all $x \in U$.

   Given $a, b \in U$ let $\varphi : \mathbb{R} \to U$ be a differentiable mapping with $\varphi(0) = a$ and $\varphi(1) = b$, which exists from the definition of connectedness. Define $g = f \circ \varphi : \mathbb{R} \to \mathbb{R}$, then

   $$g'(t) = \nabla f(\varphi(t)) \cdot \varphi'(t) = 0$$

   for all $t \in \mathbb{R}$. Therefore g is constant on $[0, 1]$ from above, so

   $$f(a) = f(\varphi(0)) = g(0) = g(1) = f(\varphi(1)) = f(b)$$

   Therefore f is constant on U.

7. *If* $T = Df(x_0) \neq I$ *show that* $\tilde{f} = T^{-1} \circ f$ *satisfies the hypotheses of the inverse mapping theorem,* $D\tilde{f}(x_0) = I$, *and the conclusion holds for* f *if it holds for* $\tilde{f}$.

Let $\tilde{f}$ be differentiable at x_0, then from the linearity of T^{-1} it follows that

$$D\tilde{f}(x_0) = D(T^{-1} \circ f)(x_0) = DT^{-1}(f(x_0)) \circ Df(a) = T^{-1} \circ T = I$$

The identity map is clearly continuous and an isomorphism, so $\tilde{f}$ is C^1 and $Df(x_0)$ is an isomorphism.

8. *Show that for any* $A \in M_n(F)$, $\|A - I\| < 1$ *implies* $A \in GL_n(F)$.