Math 125B
Vogler

Inverse Function Theorem Proof

Inverse Function Theorem

Let $V \subseteq \mathbb{R}^n$ be open and $\vec{f}: V \to \mathbb{R}^n$ with $\vec{f} \in \mathcal{C}'(V)$. If $\Delta_{\vec{f}}(\vec{a}) \neq 0$ for some $\vec{a} \in V$, then there exists an open set $W \subseteq V$ where $\vec{a} \in W$ and the following holds:

1) $\vec{f}$ is 1-1 on W, 2) $\vec{f}^{-1} \in \mathcal{C}'(\vec{f}(W))$,

3) $\forall \vec{y} \in \vec{f}(W)$, $D(\vec{f}^{-1})(\vec{y}) = [D\vec{f}(\vec{f}^{-1}(y))]^{-1}$. ← Matrix Inversion

Pf By Lemma 11.40, $\exists\, r > 0$ such that for the open ball $B := B_r(\vec{a})$, we have $\vec{f}$ is 1-1 and $\Delta_{\vec{f}} \neq 0$ on B. Also, we have

$$\Delta := \det\left[\frac{\partial f_i}{\partial x_j}(c_i)\right] \neq 0 \quad \forall c_i \in B.$$

By Lemma 11.39, $\vec{f}^{-1} \in \mathcal{C}(\vec{f}(B))$. Note $B \subseteq V$ by Lemma 11.40.

Choose $W \subseteq B \subseteq V$ to be an open ball centered at $\vec{a}$, but smaller than B (i.e. $W = B_{r_0}(\vec{a})$ with $0 < r_0 < r$). Then, $\vec{f}$ is 1-1 on W, proving claim 1). Also, $\vec{f}(W)$ must be open since f^{-1} is continuous on $\vec{f}(W)$ by Lemma 11.39.

We want to show the first partials of $\vec{f}^{-1}$ exist and are continuous on $\vec{f}(W)$. Fix $\vec{y} \in f(W)$ along with $i, k = 1, \ldots, n$. Choose $t \in \mathbb{R} - \{0\}$ small enough where $\vec{y} + t\hat{e}_k \in f(W)$. Let $\vec{x} = f^{-1}(\vec{y})$ and $\vec{w} = f^{-1}(\vec{y} + t\hat{e}_k)$. Notice for each $i = 1, \ldots, n$, we have

$$f_i(\vec{w}) - f_i(\vec{x}) = \begin{cases} t & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases} \quad \text{or} \quad \frac{f_i(\vec{w}) - f_i(\vec{x})}{t} = \begin{cases} 1 & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases}$$

Thus, by Mean Value Thm (Thm 11.30), there exists n points (one for each i) $\vec{c}_i \in L(\vec{x};\vec{w})$ where

$$\nabla f_i(\vec{c}_i) \cdot \frac{\vec{x}-\vec{w}}{t} = \frac{f_i(\vec{x})-f_i(\vec{w})}{t} = \begin{cases} 1 & \text{if } k=i \\ 0 & \text{if } k\neq i \end{cases} \quad \text{for } i=1,\dots,n \quad (1)$$

Define $\vec{z} = \frac{\vec{x}-\vec{w}}{t}$ and matrix $A = \left[\frac{\partial f_i}{\partial x_j}(c_i)\right]_{n\times n}$ so (1) becomes

$$A\vec{z} = \begin{array}{cc} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} & \begin{array}{l} \\ \\ \leftarrow k\text{th slot} \\ \\ \\ \end{array} \end{array} =: \vec{k} \quad (2).$$

Equation (2) is a system of n linear equations and n unknowns (i.e $\frac{x_j - w_j}{t}$) where the coefficient matrix A has determinant $\Delta \neq 0$ on $W \subseteq B$ (by our choice of B). Using Cramer's Rule (see Appendix C), solutions to (2) satisfy

$$\frac{(f^{-1})_j(\vec{y}+t\hat{e}_k) - (f^{-1})_j(\vec{y})}{t} := \frac{x_j - w_j}{t} = Q_j(t) := \frac{\det(A(j))}{\Delta}, \quad (3)$$

where $A(j)$ is the matrix formed by replacing the jth column of A with the vector $\vec{k}$. Notice $Q_j(t)$ is the quotient of determinants for matrices which have entries of $0, 1,$ or first-order partials evaluated at the $\vec{c}_i$'s. When $t \to 0$, we have $\vec{w} \to \vec{x}$, $\vec{c}_i \to \vec{x}$ $\forall i=1,\dots,n$, and $\vec{y}+t\hat{e}_k \to \vec{y}$. Also, $Q_j(t) \to Q_j$, which is a quotient of determinants whose entries consist of $0, 1,$ or first-order partials evaluated at $\vec{x} = f^{-1}(\vec{y})$.

Since $\vec{f}^{-1}$ is continuous on $\vec{f}(W)$, Q_j must be continuous $\forall \vec{y} \in \vec{f}(W)$ through algebraic and functional composition of continuous functions. Taking the limit of (3) as $t \to 0$, we have

$$\frac{\partial (f^{-1})_j}{\partial x_k}(\vec{y}) = Q_j \quad \forall j=1,\dots,n, \ \forall y \in \vec{f}(W). \quad (4)$$

Since $k=1,\dots,n$ was arbitrary, (4) holds for all $k=1,\dots,n$. Therefore, all 1st-order partials of f^{-1} exist on $\vec{f}(W)$, and we conclude $\vec{f}^{-1} \in \mathcal{C}'(\vec{f}(W))$, proving claim 2).

For claim 3), fix $\vec{y} \in \vec{f}(W)$. By HW 11.2.8 and Chain Rule, we can directly compute

$$I = D[I\vec{y}] = D[(\vec{f} \circ \vec{f}^{-1})(\vec{y})] = D\vec{f}(\vec{f}^{-1}(\vec{y})) \cdot D\vec{f}^{-1}(\vec{y}),$$

where I is the $n \times n$ identity matrix. By the uniqueness of matrix inverses, we conclude

$$D\vec{f}^{-1}(\vec{y}) = [D\vec{f}(\vec{f}^{-1}(\vec{y}))]^{-1},$$

proving claim 3). $\blacksquare$