Economics 204

Lecture 12–Tuesday, August 11, 2009

Inverse and Implicit Function Theorems, and Generic Methods:

Section 4.3 (Conclusion), Regular and Critical Points and Values:

Definition 1 Suppose $X \subseteq \mathbf{R}^n$ is open. Suppose $f : X \to \mathbf{R}^m$ is differentiable at $x \in X$, and let $W = \{e_1, \ldots, e_n\}$ denote the standard basis of $\mathbf{R}^n$. Then $df_x \in L(\mathbf{R}^n, \mathbf{R}^m)$, and

$$\begin{aligned}
\operatorname{Rank} df_x &= \dim \operatorname{Im}(df) \\
&= \dim \operatorname{span}\{df_x(e_1), \ldots, df_x(e_n)\} \\
&= \dim \operatorname{span}\{Df(x)e_1, \ldots, Df(x)e_n\} \\
&= \dim \operatorname{span}\{\text{column 1 of } Df(x), \ldots, \text{column n of} Df(x)\} \\
&= \operatorname{Rank} Df(x)
\end{aligned}$$

Thus,

$$\operatorname{Rank}(df_x) \leq \min\{m, n\}$$

We say

- x is a *regular point* of f if $\operatorname{Rank}(df_x) = \min\{m, n\}$.
- x is a *critical point* of f if $\operatorname{Rank}(df_x) < \min\{m, n\}$.
- y is a *critical value* of f if there exists $x \in X$, $f(x) = y$, x is a critical point of f.
- y is a *regular value* of f if y is not a critical value of f (notice this has the counterintuitive implication that if $y \notin f(X)$, then y is automatically a regular value of f).

Remark: The definition of regular point and critical point in de la Fuente (as well as in Mas-Colell, Whinston, and Green) is different: they use m rather than $\min\{m, n\}$. I think the definition I have given is more natural. If $m \leq n$, the two are equivalent. If $m > n$, then since $\text{Rank}\,(df_x) \leq \min\{m, n\}$, then every $x \in X$ will be a critical point in the de la Fuente and MWG definitions, and every $y \in f(X)$ will be a critical value. In the definition I have given, a point is critical if the rank is smaller than the largest it could possibly be. The two important theorems (Sard's Theorem and the Transversality Theorem) concerning critical values are true with either definition.

Example: Consider the function $f : (0, 2\pi) \to \mathbf{R}$ defined by

$$f(x) = \sin x$$

Then $f'(x) = \cos x$, so $f'(x) = 0$ for $x = \pi/2$ and $x = 3\pi/2$. $Df(x)$ is the 1×1 matrix $(f'(x))$, so $\text{Rank}\, df_x = \text{Rank}\, Df(x) = 1$ if and only if $f'(x) \neq 0$. Thus, the critical points of f are $\pi/2$ and $3\pi/2$, so the set of regular points of f is

$$(0, \pi/2) \cup (\pi/2, 3\pi/2) \cup (3\pi/2, 2\pi)$$

The critical values of f are $f(\pi/2) = \sin(\pi/2) = 1$ and $f(3\pi/2) = \sin(3\pi/2) = -1$; the set of regular values of f is

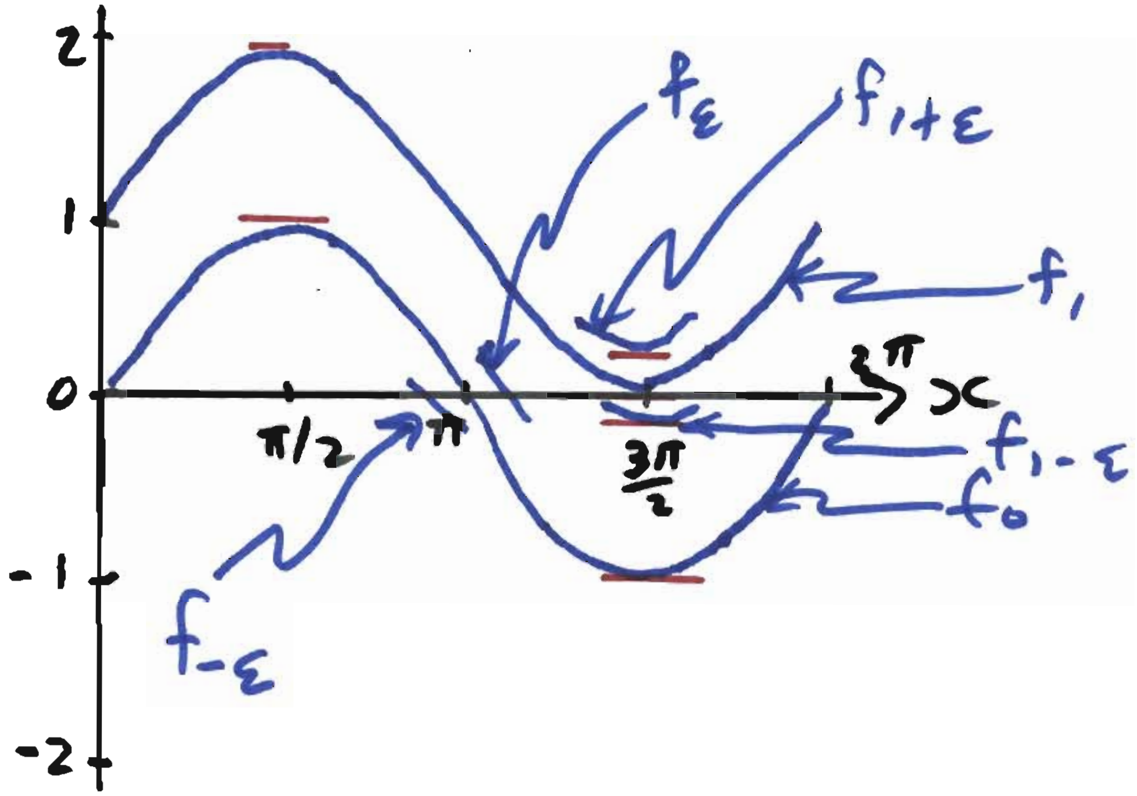
$$(-\infty, -1) \cup (-1, 1) \cup (1, \infty)$$

Notice that 0 is not a critical value. Given $\alpha \in \mathbf{R}$, consider the perturbed function

$$f_\alpha(x) = f(x) + \alpha$$

Notice that $f'_\alpha(x) = f'(x)$, so the critical points of f_α are the same as those of f. For α close to zero, the solution to the equation

$$f_\alpha(x) = 0$$



$$f_\alpha(x) = \sin x + \alpha$$

near $x = \pi$ moves smoothly with respect to changes in α; the direction a solution moves is determined by the sign of f'_α.

Now, let $\alpha = 1$.

$$\begin{aligned} f_1(x) = 0 \quad &\Leftrightarrow \quad \sin x + 1 = 0 \\ &\Leftrightarrow \quad \sin x = -1 \\ &\Leftrightarrow \quad x = \frac{3\pi}{2} \end{aligned}$$

Since $3\pi/2$ is a critical point of f_1, 0 is a critical value of f_1.

Consider the correspondence

$$\Psi(\alpha) = \{x : f_\alpha(x) = 0\} \text{ for } \alpha \in [0, 2]$$

(De la Fuente requires that correspondences be nonempty-valued, but we don't; we shall see that $\Psi(\alpha) = \emptyset$ for $\alpha > 1$.) Note that for α close to one, we have the following:

- if $\alpha = 1$, the equation $f_\alpha(x) = 0$ has one solution, $3\pi/2$, so $\Psi(1) = \{3\pi/2\}$.
- if $\alpha < 1$, the equation $f_\alpha(x) = 0$ has two solutions, both near $3\pi/2$.
- if $\alpha > 1$, the equation $f_\alpha(x) = 0$ has no solutions; the unique solution for $\alpha = 1$ disappears in a puff of smoke. Hence Ψ is not lower hemicontinuous at $\alpha = 1$. Ψ is lower hemicontinuous at all other $\alpha \in [0, 2]$.
- Let $\hat{\Psi}$ be the restriction of Ψ to the domain $[0, 1]$. Then $\hat{\Psi}$ is lower hemicontinuous at 1, but because the unique element of $\Psi(1)$ splits into two points that move in opposite directions as α decreases, we cannot make sense of comparative statics questions such as "in what direction does the solution to $f_\alpha(x) = 0$ move if α starts at 1 and is decreased?"

Thus, if 0 is a critical value of a function f, then the solutions to the equation $f(x) = 0$ may behave badly in response to small perturbations of f; we will return to this in Lecture 13.

Inverse Function Theorem:

Theorem 2 (4.6, Inverse Function Theorem) *Suppose*

$X \subseteq \mathbf{R}^n$ *is open,* $f : X \to \mathbf{R}^n$*,* $f \in C^1(X)$*,* $x_0 \in X$*. If*

$$\det\left(Df(x_0)\right) \neq 0$$

(i.e. x_0 *is a regular point of* f*) then there are open neighborhoods* U *of* x_0 *and* V *of* $f(x_0)$ *such that*

$$\begin{aligned}
f : U \to V \quad & \textit{is} \quad \textit{one-to-one and onto} \\
f^{-1} : V \to U \quad & \textit{is} \quad C^1 \\
\left(D\left(f^{-1}\right)\right)(f(x_0)) \quad & = \quad (Df(x_0))^{-1} \\
f \in C^n \quad & \Rightarrow \quad f^{-1} \in C^n
\end{aligned}$$

Remark: f is one-to-one only on U; it need not be one-to-one globally. f^{-1} is only a local inverse. To see the formula for $D(f^{-1})$, let Id_U denote the identity function from U to U and I the $n \times n$ identity matrix. Then

$$\begin{aligned}
\left(D\left(f^{-1}\right)(f(x_0))\right) Df(x_0) &= D\left(f^{-1} \circ f\right)(x_0) \\
&= D\left(Id_U\right)(x_0) \\
&= I \\
\left(D\left(f^{-1}\right)(f(x_0))\right) Df(x_0) &= (Df(x_0))^{-1}
\end{aligned}$$

Proof: Read the proof in de la Fuente. This is pretty hard. The idea is that since $\det Df(x_0) \neq 0$, then $df_{x_0} : \mathbf{R}^n \to \mathbf{R}^n$ is one-to-one and onto. You need to find a neighborhood U of x_0 sufficiently small such that the Contraction Mapping Theorem implies that f is one-to-one and onto. ■

Read Section 4.5 on Your Own

Section 5.2, Implicit Function Theorem

Function $F(x, \omega)$; x is a variable vector (e.g. price vector); ω is a vector of parameters (e.g. endowments).

x is determined implicitly as a function of ω by the equation

$$F(x(\omega), \omega) = 0$$

e.g. Walrasian equilibrium (market-clearing) prices determined as an implicit function of endowments.

$$D_x F(x, \omega) \text{denotes the matrix of}$$

$$\text{partial derivatives with respect to } x \text{ only}$$

Theorem 3 (2.2, Implicit Function Theorem) *Suppose*

$X \subseteq \mathbf{R}^n$ and $\Omega \subseteq \mathbf{R}^p$ are open and $F : X \times \Omega \to \mathbf{R}^n$ is C^1. Suppose

$$\begin{aligned} F(x_0, \omega_0) &= 0 \\ \det(D_x F(x_0, \omega_0)) &\neq 0 \end{aligned}$$

i.e. x_0 is a regular point of $F(\cdot, \omega_0)$. Then there are open neighborhoods U of x_0 ($U \subseteq X$) and W of ω_0 such that

$$\forall_{\omega \in W} \exists!_{x \in U} \; F(x, \omega) = 0$$

Let $g(\omega)$ be that unique x. Then

$$\begin{aligned} g : W \to X \quad & is \quad C^1 \\ Dg(\omega_0) &= -\left[D_x F(x_0, \omega_0)\right]^{-1} \left[D_\omega F(x_0, \omega_0)\right] \\ F \in C^k &\Rightarrow g \in C^k \end{aligned}$$

If 0 is a regular value of $F(\cdot, \omega_0)$, then the correspondence

$$\omega \to \{x : F(x, \omega) = 0\}$$

is lower hemicontinuous at ω_0.

Proof: Use the Inverse Function Theorem in the right way. Why is the formula for Dg correct? Assuming the implicit function exists and is differentiable,

$$\begin{aligned} 0 &= DF(g(\omega), \omega)(\omega_0) \\ &= D_x F(x_0, \omega_0) Dg(\omega_0) + D_\omega F(x_0, \omega_0) \\ Dg(\omega_0) &= -\left(D_x F(x_0, \omega_0)\right)^{-1} D_\omega F(x_0, \omega_0) \end{aligned}$$

The following argument outlines the proof that g is differentiable:

$$\begin{aligned} F(x_0, \omega_0 + h) &= F(x_0, \omega_0) + D_\omega F(x_0, \omega_0) h + o(h) \\ &= D_\omega F(x_0, \omega_0) h + o(h) \end{aligned}$$

Solve for Δx that brings F back to zero:

$$\begin{aligned} 0 &= F(x_0 + \Delta x, \omega_0 + h) \\ &= F(x_0, \omega_0 + h) + D_x F(x_0, \omega_0 + h)\Delta x + o(\Delta x) \\ &= F(x_0, \omega_0) + D_\omega F(x_0, \omega_0) h + D_x F(x_0, \omega_0 + h)\Delta x \\ &\quad + o(\Delta x) + o(h) \\ &= D_\omega f(x_0, \omega_0) h + D_x F(x_0, \omega_0 + h)\Delta x + o(\Delta x) + o(h) \\ &D_x F(x_0, \omega_0 + h)\Delta x \\ &= -D_\omega F(x_0, \omega_0) h + o(\Delta x) + o(h) \end{aligned}$$

Because F is C^1 and the determinant is a continuous functions of the entries of the matrix, we have $\det D_x F(x_0, \omega_0 + h) \neq 0$ for h sufficiently small, so

$$\begin{aligned} \Delta x &= -\left[D_x F(x_0, \omega_0 + h)\right]^{-1} D_\omega F(x_0, \omega_0) h \\ &\quad + o(\Delta x) + o(h) \\ &= -\left[D_x F(x_0, \omega_0) + o(1)\right]^{-1} D_\omega F(x_0, \omega_0) h \\ &\quad + o(\Delta x) + o(h) \text{ since } F \in C^1 \end{aligned}$$

$$
\begin{aligned}
&= -\left[D_x F(x_0,\omega_0)\right]^{-1} D_\omega F(x_0,\omega_0)h \\
&\quad +o(\Delta x)+o(h) \text{ since } F \in C^1 \\
|\Delta x + o(\Delta x)\| &= O(h) \\
\Rightarrow |\Delta x\| &= O(h) \\
\Rightarrow o(\Delta x) &= o(h) \\
\Rightarrow \Delta x &= -\left[D_x F(x_0,\omega_0)\right]^{-1} D_\omega F(x_0,\omega_0)h + o(h)
\end{aligned}
$$

By the definition of the derivative,

$$Dg(\omega_0) = -\left[D_x F(x_0,\omega_0)\right]^{-1} D_\omega F(x_0,\omega_0)$$

If 0 is a regular value of $F(\cdot,\omega_0)=0$, then given any $x_0 \in \{x : F(x,\omega_0)=0\}$, we can find a local implicit function g; in other words, if ω is sufficiently close to ω_0, then $g(\omega) \in \{x : F(x,\omega)=0\}$; the continuity of g then shows that the correspondence $\{x : F(x,\omega)=0\}$ is lower hemicontinuous at ω_0. ■

Transversality and Genericity

Definition 4 Suppose $A \subseteq \mathbf{R}^n$. A has *Lebesgue measure zero* if, for every $\varepsilon > 0$, there is a countable collection of rectangles $I_1, I_2, \ldots$ such that

$$\sum_{k=1}^{\infty} \mathrm{Vol}\,(I_k) < \varepsilon \text{ and } A \subseteq \cup_{k=1}^{\infty} I_k$$

Notice that this defines Lebesgue measure zero without defining Lebesgue measure(!)

This is a natural formulation of the notion that A is a small set:

"If you choose $x \in \mathbf{R}^n$ at random,

the probability that $x \in A$ is zero."

It is easy to show that

$$
\begin{aligned}
&A_n \text{ has Lebesgue measure zero} \\
&\Rightarrow \ \cup_{n\in\mathbf{N}} A_n \text{ has Lebesgue measure zero}
\end{aligned}
$$

In particular, $\mathbf{Q}$ and every countable set has Lebesgue measure zero.

A function may have many critical points; for example, if a function is constant on an interval, then every element of the interval is a critical point. But it can't have many critical *values.*

Theorem 5 (2.4, Sard's Theorem) *Let $X \subseteq \mathbf{R}^n$ be open, $f : X \to \mathbf{R}^m$, f is C^r with $r \geq 1+\max\{0, n-m\}$. Then the set of all critical values of f has Lebesgue measure zero.*

Proof: First, we give a false proof that conveys the essential idea as to why the theorem is true; it can be turned into a correct proof. Suppose $m = n$. Let C be the set of critical points of f, V the set of critical values. Then

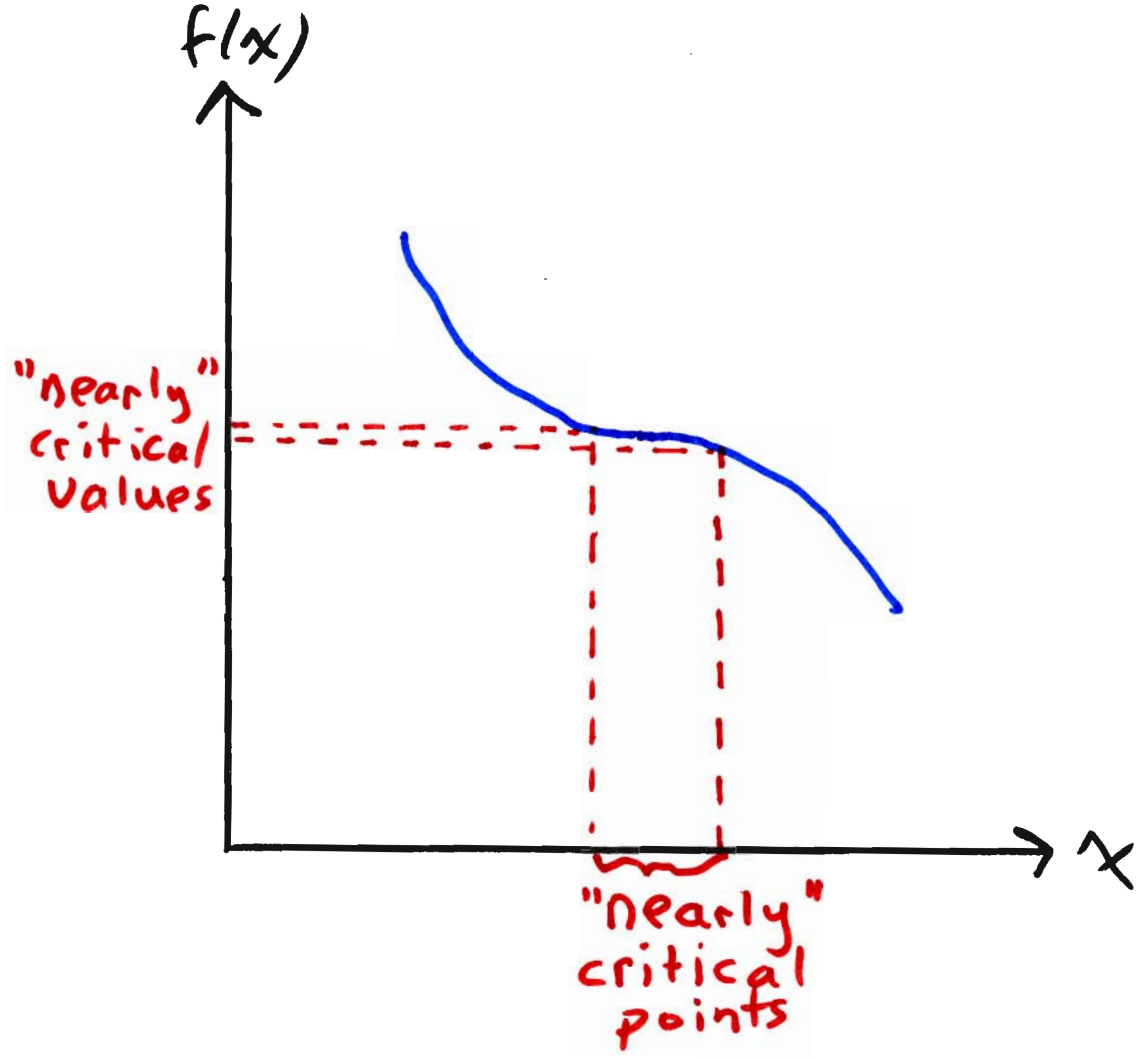
$$\begin{aligned}
\mathrm{Vol}\,(V) &= \mathrm{Vol}\,(f(C)) \\
&\leq \int_C |\det Df(x)|\,dx \text{ (equality if } f \text{ is one-to-one)} \\
&= \int_C 0\,dx \\
&= 0
\end{aligned}$$

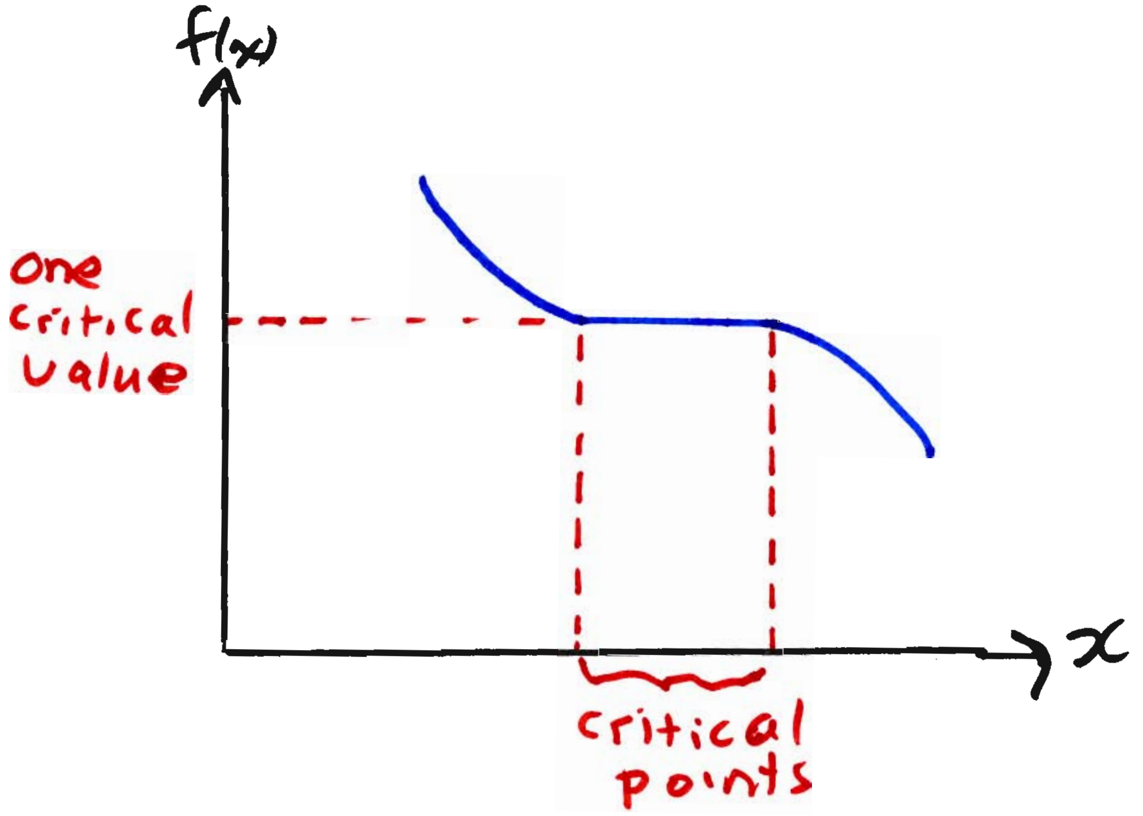
Now, we outline how to turn this into a proof. First, show that we can write $X = \cup_{j\in\mathbf{N}} X_j$, where each X_j is a compact subset of $[-j, j]^n$. Let $C_j = C \cap X_j$. Fix j for now. Since f is C^1,

$$\begin{aligned}
x_k \to x \quad &\Rightarrow \quad \det Df(x_k) \to \det Df(x) \\
\{x_k\} \subseteq C_j, x_k \to x \quad &\Rightarrow \quad \det Df(x) = 0 \Rightarrow x \in C_j
\end{aligned}$$

so C_j is closed, hence compact. Since X is open and C_j is compact, there exists $\delta_1 > 0$ such that

$$B_{\delta_1}[C_j] = \cup_{x\in C_j} B_{\delta_1}[x] \subseteq X$$


f(x)
"nearly" critical values
x
"nearly" critical points


f(x)
one critical value
x
critical points

$B_{\delta_1}[C_j]$ is bounded, and, using the compactness of C_j, one can show it is closed, so it is compact. $\det Df(x)$ is continuous on $B_{\delta_1}[C_j]$, so it is uniformly continuous on $B_{\delta_1}[C_j]$, so given $\varepsilon > 0$, we can find $\delta \leq \delta_1$ such that $B_\delta[C_j] \subseteq [-2j, 2j]^n$ and

$$x \in B_\delta[C_j] \Rightarrow \det |Df(x)| \leq \frac{\varepsilon}{2 \cdot 4^n j^n}$$

Then

$$\begin{aligned} f(C_j) &\subseteq f(B_\delta[C_j]) \\ \mathrm{Vol}\,(f(B_\delta[C_j])) &\leq \int_{[-2j,2j]^n} \frac{\varepsilon}{2 \cdot 4^n j^n}\, dx \\ &= \frac{\varepsilon}{2} \end{aligned}$$

Since f is C^1, show that $f(C_j)$ can be covered by a countable collection of rectangles of total volume less than ε. Since $\varepsilon > 0$ is arbitrary, $f(C_j)$ has Lebesgue measure zero. Then

$$f(C) = f\left(\cup_{j \in \mathbf{N}} C_j\right) = \cup_{n \in \mathbf{N}} f(C_j)$$

is a countable union of sets of Lebesgue measure zero, so $f(C)$ has Lebesgue measure zero. ■

Significance of Sard's Theorem:

- Given a randomly chosen function f, it is very unlikely that zero will be a critical value of f.
- If by some fluke zero *is* a critical value of f, then a slight perturbation of f will make zero a regular value.
- If zero is a regular value of f, we can apply the Inverse Function Theorem or the Implicit Function Theorem, as appropriate given the dimensions of the domain and range of f.

Recall that our definition of critical point differed from de la Fuente's in the case $m > n$:

- If $m > n$, then every $x \in X$ is critical using de la Fuente's definition, because

$$\operatorname{Rank} Df(x) \leq n < m$$

- Thus, every $y \in f(X)$ is a critical value, using de la Fuente's definition.

- This does not contradict Sard's Theorem, since one can show that $f(X)$ is a set of Lebesgue measure zero when $m > n$ and $f \in C^1$.