Recall: • $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at x if $\exists\, Df(x) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ s.t.
$$f(x+v) = f(x) + Df(x)v + o(|v|).$$
- $f \in C^1(U, \mathbb{R}^m)$ if everywhere differentiable and $Df: U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous.
- as a matrix, entries in $Df(x)$ are partial derivatives $\partial f_i / \partial x_j$.
- operator norm: $\|Df(x)\| = \sup_{v \neq 0} \dfrac{|Df(x)v|}{|v|}$

- Higher order derivatives: f is C^2 if $Df: U \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) \simeq \mathbb{R}^{n \times m}$ is C^1, etc.

The main important fact about higher partial derivatives is:

Prop: if $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}\right)$ exist and are continuous then $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{\partial^2 f}{\partial x_j \partial x_i}$

Pf: enough to consider the case of $f(x,y)$. For h and k small $\neq 0$, consider
$$\frac{1}{hk}\Big(f(x+h, y+k) - f(x+h, y) - f(x, y+k) + f(x,y)\Big)$$
writing this in terms of $g(x,y) = \dfrac{f(x, y+k) - f(x,y)}{k}$, this is $\dfrac{1}{h}\big(g(x+h, y) - g(x,y)\big)$

so by mean value thm for $\dfrac{\partial g}{\partial x}$, $\exists\, h_1 \in (0,h)$ st. this equals
$$\frac{\partial g}{\partial x}(x+h_1, y) = \frac{1}{k}\left(\frac{\partial f}{\partial x}(x+h_1, y+k) - \frac{\partial f}{\partial x}(x+h_1, y)\right).$$
In turn, by mean value thm for $\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right)$, $\exists\, k_1 \in (0,k)$ st. this equals $\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right)(x+h_1, y+k_1)$.

Doing the same calculation in opposite order shows $= \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right)(x+h_2, y+k_2)$ for some $h_2 \in (0,h)$, $k_2 \in (0,k)$.

Since these 2nd derivatives are continuous by assumption, taking limits as $h, k \to 0$ gives the result. $\square$

- Hence: the Hessian matrix $H = \left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}\right)$ is symmetric. and should be interpreted as a symmetric bilinear form on tangent vectors. If $f \in C^2$ then
$$f(x+v) = f(x) + Df(x)\cdot v + \frac{1}{2} H(x)(v,v) + o(|v|^2) \qquad \text{(\& so on, Taylor!)}.$$

- Because of the local approximation $f(x+v) = f(x) + Df(x)v + r(v)$, the behavior of $Df(x)$ governs that of f near x. In particular:

  → if $Df(x)$ is injective then f is injective on a (suff. small) neighborhood of x.

  → if $Df(x)$ is surjective then f maps a neighborhood of x surjectively onto a nbd of $f(x)$.

When both hold, f is a local diffeomorphism, by the inverse function theorem.

Def: a map $f: U \to V$ between open subsets of $\mathbb{R}^n$ is a diffeomorphism if it is a homeomorphism and both f and f^{-1} are C^1.

Thm: Let $p \in E \subset \mathbb{R}^n$ open, $f: E \to \mathbb{R}^n$ C^1, suppose $Df(p): \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism (ie. $\det Df(p) \neq 0$). Then f is a local diffeomorphism at p, ie. $\exists U \ni p$ neighborhood st. f is a diffeomorphism between $U \subset E$ and $f(U) \subset \mathbb{R}^n$.

The proof uses two main ingredients:

1. mean value inequality: $\sup \|Df\| \leq M \Rightarrow |f(b) - f(a)| \leq M|b-a|$.
2. contraction mapping principle: X complete metric space, $\varphi: X \to X$ contraction $(d(\varphi(x), \varphi(y)) \leq \alpha d(x,y)$ for some $\alpha < 1) \Rightarrow \varphi$ has a unique fixed point.

(__Proof__: • existence: let $x_0 \in X$, set $x_{n+1} = \varphi(x_n)$, then $d(x_n, x_{n+1}) \leq \alpha\, d(x_{n-1}, x_n)$ so $d(x_n, x_{n+1}) \leq \alpha^n d(x_0, x_1)$ so (x_n) is Cauchy, hence converges to some $x \in X$.
Moreover $x_{n+1} = \varphi(x_n) \longrightarrow \varphi(x)$, but $\lim\limits_{n\to\infty} x_{n+1} = \lim\limits_{n\to\infty} x_n = x$, hence $\varphi(x) = x$.
• uniqueness: if $\varphi(x) = x$ and $\varphi(y) = y$ then $d(\varphi(x), \varphi(y)) = d(x,y) \leq \alpha d(x,y)$, so $x = y$.)

__Pf. of inverse function theorem:__

- After a linear change of variables, we can assume $p = 0$, $f(0) = 0$, $Df(0) = Id$.
- Since $f \in C^1$, Df is continuous, so $\exists$ ball $B_r(0)$ st. $\|Df(x) - I\| \leq \frac{1}{2}$ for $|x| \leq r$
- Now, given $y_0 \in \mathbb{R}^n$, let $\varphi(x) = x + (y_0 - f(x))$.
  "next guess using Newton's method to find x_0 st. $f(x_0) = y_0$, given $f(x)$, using $Df \sim I$."

__Key obs^n__: $\varphi(x) = x$ iff $f(x) = y_0$, and for $|x| \leq r$ we have $\|D\varphi(x)\| = \|I - Df(x)\| \leq \frac{1}{2}$.

- Assume $|y_0| < \frac{r}{2}$. Since $\varphi(0) = y_0$ and $\|D\varphi\| \leq \frac{1}{2}$ for $|x| \leq r$, the mean value inequality gives for $|x_i| \leq r$, $|\varphi(x_1) - \varphi(x_2)| \leq \frac{1}{2}|x_1 - x_2|$.
  and also $|\varphi(x)| \leq |y_0| + \frac{|x|}{2} < r$. $(*)$

  (by $(*)$, $|x_0| = |\varphi(x_0)| < r$.)

So φ is a contracting map from $\overline{B_r(0)}$ to itself, hence $\exists!$ fixed point $x_0 \in B_r(0)$
Thus $\forall y_0 \in B_{\frac{r}{2}}(0)$, $\exists!\, x_0 \in B_r(0)$ st. $f(x_0) = y_0$. $(**)$

- Now let $V = B_{\frac{r}{2}}(0)$, $U = f^{-1}(V) \cap B_r(0)$, then U, V are open (f continuous) and $f_{|U}: U \to V$ is a bijection by $(**)$. Let $g: V \to U$ the inverse map.
- __Claim:__ g is differentiable and $Dg(y) = Df(x)^{-1}$ where $x = g(y)$ $(y = f(x))$

  __Pf:__ fix $y_0 \in V$, $x_0 = g(y_0) \in U$, let $\varphi(x) = x + (y_0 - f(x))$ as above, with $\varphi(x_0) = x_0$.
  For $w \in \mathbb{R}^n$ small (so $|y_0 + w| < \frac{r}{2}$), write $g(y_0 + w) = x_0 + v$, so $f(x_0 + v) = y_0 + w$.
  Then $\varphi(x_0 + v) = (x_0 + v) + (y_0 - (y_0 + w)) = x_0 + v - w$, vs. $\varphi(x_0) = x_0$.
  But we've shown φ is contracting, $|\varphi(x_0 + v) - \varphi(x_0)| = |v - w| \leq \frac{1}{2}|v|$.
  Hence $|w| \geq \frac{1}{2}|v|$ by triangle inequality, ie. $|v| \leq 2|w|$.

Given $\varepsilon>0$ $\exists\,\delta$ st. $|v|<\delta \Rightarrow \left|f(x_0+v)-f(x_0)-Df(x_0)v\right| < \frac{\varepsilon}{2}|v|$.

$\Rightarrow$ for $|w|<\frac{\delta}{2}$, $\left|(y_0+w)-y_0-Df(x_0)v\right| < \frac{\varepsilon}{2}|v| \leq \varepsilon|w|$.

Applying $Df(x_0)^{-1}$: for $|w|<\frac{\delta}{2}$, $|Df(x_0)^{-1}w - v| \leq \|Df(x_0)^{-1}\|\ |w-Df(x_0)v|$
$< \varepsilon\ \|Df(x_0)^{-1}\|\ |w|$.

Recalling $v=g(y_0+w)-g(y_0)$, this yields
$$g(y_0+w) = g(y_0) + Df(x_0)^{-1}w + o(|w|). \qquad \square.$$

- the continuity of $Dg = Df^{-1}\circ g$ then follows from the continuity of Df and of g itself. $\square$

---

* Implicit function theorem:

$\mathbb{R}^n\times\mathbb{R}^m \supset E$ open, $f: E\to\mathbb{R}^m$ differentiable.
$(x,y)\mapsto f(x,y)$

Write $Df(x,y): \mathbb{R}^n\oplus\mathbb{R}^m \to \mathbb{R}^m$ as $Df_x \oplus Df_y$, $Df_x: \mathbb{R}^n\to\mathbb{R}^m$ first n variables, $Df_y: \mathbb{R}^m\to\mathbb{R}^m$ last m variables

Assume $f(x_0,y_0)=0$ and Df_y is invertible ($\det Df_y \neq 0$) at $(x_0,y_0)\in E$

Then $\exists\ U\ni x_0,\ V\ni y_0$ open st. $\forall x\in U\ \exists!\ y =: g(x)\in V$ st. $f(x,y)=0$.

Moreover, $g: U\to V$ defined by $f(x,g(x))=0\ \forall x\in U$ is differentiable, and $Dg = -(Df_y)^{-1}Df_x$

This follows from the inverse function theorem by considering
$$F: \mathbb{R}^{n+m}\supset E \longrightarrow \mathbb{R}^{n+m} \qquad DF(x_0,y_0) = \left(\begin{array}{c|c} I & 0 \\ \hline Df_x & Df_y\end{array}\right) \text{ invertible } \checkmark$$
$$F(x,y) = (x, f(x,y)).$$

This has an inverse G over a nbd. of $F(x_0,y_0)=(x_0,0)$.

Near (x_0,y_0), $f(x,y)=0 \Leftrightarrow F(x,y)=(x,0) \Leftrightarrow (x,y)=G(x,0)$.

So we let $g(x)=$ second component of $G(x,0)$. $\square$

* Given a differentiable $f: \mathbb{R}^{n+m}\to\mathbb{R}^m$, and a point at which Df is surjective, we can always find a subset of coordinates $(x_i)_{i\in I}$ ($I\subset\{1..m+n\}, |I|=m$) st. the corresponding part of Df is invertible $\Rightarrow$ can apply implicit f$^{\underline{n}}$ theorem to describe the zero set of f by eq$^{\underline{ns}}$ $(x_i)_{i\in I} = g(x_j, j\notin I)$.

In particular, a <u>hypersurface</u> $S\subset\mathbb{R}^n$ = closed subset which is locally the zero set of a differentiable real-valued function f with $Df\neq 0$. Using implicit f$^{\underline{n}}$ theorem, S can be locally described as the graph $x_j = g(x_i, i\neq j)$ of some diff$^{\underline{able}}$ $g: \mathbb{R}^{n-1}\to\mathbb{R}$.

Eg. a diff$^{\underline{able}}$ curve in $\mathbb{R}^2$ is locally a graph $x=f(y)$ or $y=f(x)$.

Iterated and Riemann integrals in several variables

* f continuous function on an n-cell $I = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$

$\Rightarrow$ we can define $\int_I f = \int_I f \, dx_1 \dots dx_n = \int_I f \, |dx|$ ← why? clearer after diff. forms

either 1) as iterated integral:

$$\int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \dots \left(\int_{a_n}^{b_n} f(x_1 \dots x_n) \, dx_n \right) \dots dx_2 \right) dx_1 \qquad \text{or any order}$$

2) as Riemann integral: split I into small cubes Q_i, and bound f between piecewise constant functions $s = s_i = \min f(Q_i)$ on $\mathrm{int}(Q_i)$
$S = S_i = \max f(Q_i)$ — " —

$\rightarrow$ estimate $\sum s_i \, \mathrm{vol}(Q_i) \le \int_I f \, |dx| \le \sum S_i \, \mathrm{vol}(Q_i)$

If f is continuous, hence uniformly continuous, then $\sup |S - s| \to 0$ as $\mathrm{diam}(Q_i) \to 0$, so this defines the integral uniquely.

Fubini's thm says: for continuous f, the iterated integrals for different orders of integration are all equal.

* if f is only piecewise continuous, integrability still holds if the regions of I where f is continuous are sufficiently regular - eg. delimited by smooth hypersurfaces. Specifically: when decomposing D into small cubes Q_i, want $\sum_{f_{|Q_i} \text{ not } C^0} \mathrm{vol}(Q_i) \longrightarrow 0$ as one subdivides further – over such cubes, $(S_i - s_i)$ doesn't $\to 0$ as step size $\to 0$, but if vol $\to 0$ we still have $\int_D (S - s) \, |dx| = \sum (S_i - s_i) \, \mathrm{vol}(Q_i) \to 0$.

* Thus we can define integrals over regions of $\mathbb{R}^n$ delimited by hypersurfaces by either
  - extending f by 0 outside of the given region, and integrating the resulting piecewise continuous function
  - using changes of coords. (via implicit function thm) to make the region of integration an n-cell. This requires change of variables ...

Thm: $\varphi : U \to V$ diffeomorphism, f continuous on V, then
$$\int_V f(y) \, |dy| = \int_U f(\varphi(x)) \, |\det D\varphi(x)| \, d|x|.$$

(won't prove. The geometric input is that if Q_i is a small cube $\ni x$ then $\varphi(Q_i) \approx$ small parallelepiped $\ni \varphi(x)$, with $\mathrm{vol}(\varphi(Q_i)) \sim |\det D\varphi(x)| \cdot \mathrm{vol}(Q_i)$.)