Multivariate Analysis

Iannis Petridis

Autum 2011

Contents

1 Mulivarible Calculus

1.1 Notation

$X \in \mathbb{R}^n$, $X = \{x_1, x_2, \dots, x_n\}$ where $x_i \in \mathbb{R}$ $\mathbb{R}^n$ is a vector space
length norm$|x| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$
If $Y, X \in \mathbb{R}^n$ and $Y = \{y_1, y_2, \dots, y_n\}$ then $X \cdot Y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$
Standard Basis:

$$e_j = (0, \dots, 0, 1, 0, \dots)$$

j-1, j, j+1

Properties of norm

$$|x| \geq 0$$

$$|x| = 0 \Leftrightarrow x = \vec{0}$$

$$|\lambda x| = |\lambda| \cdot |x|, \quad x \in \mathbb{R}^n, \quad \lambda \in \mathbb{R}$$

linear Transformation

$$T : \mathbb{R}^n \to \mathbb{R}^n$$

(i) $T(x + y) = T(x) + T(y)$

(ii) $T(\lambda x) = \lambda T(x)$

Matrix Representation of T with respect to the standard basis:

$$T(e_i) = \sum_{j=1}^{m} a_{i,j} e_j \text{ where } [T]_\epsilon^\epsilon = A = (a_i, j)_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

Given: $T : \mathbb{R}^n \to \mathbb{R}^m, S : \mathbb{R}^n \to \mathbb{R}^m$ and $U : \mathbb{R}^m \to \mathbb{R}^k$

(i) $[UT]_{kxm} = [U]_{kxm} [T]_{mxn}$

(ii) $[T + S] = [T] + [S]$

(iii) $\lambda [T] = [\lambda T]$

$T : \mathbb{R}^n \to \mathbb{R}^m, X \in \mathbb{R}^n, Y \in \mathbb{R}^m, X = (x^1, \dots, x^n), Y = (y^1, \dots, y^m)$

$$\begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix} = [T] \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$$

1.2 Functions & Continuity

$f : \mathbb{R}^n \to \mathbb{R}^m$ vector valued function
$f : A \to \mathbb{R}^m$ where $A \subset \mathbb{R}^n$
then f has components which are scalar fields.
$f^i : A \to \mathbb{R}$

$$f(x) = (f^1(x), \ldots, f^m(x))$$

$\Pi^i : \mathbb{R}^m \to \mathbb{R}$

$$\Pi^i((x)^1, \ldots, (x)^m)$$

Π^i is a linear transformation for i=1,...,m

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "f"] \arrow[dr, "f^i"'] & \mathbb{R}^m \arrow[d, "\Pi^i"] \\
& \mathbb{R}
\end{tikzcd}$$

Definition 1.1. *$f : \mathbb{R}^n \to \mathbb{R}^m$ then $\lim_{x\to a}(f(x)) = b$ means:*

$$\forall \epsilon > 0, \exists \delta > 0 \ st, \ 0 < |x - a| < \delta \Rightarrow |f(x) - b| < \epsilon$$

Definition 1.2. *f is called continuious at a if:*

$$\lim_{x\to a}(f(x)) = f(a)$$

f is called continuious at the set of A if it is continuious at a $\forall a \in Al$

Theorem 1.1 (Combination Theorm)**.** *Assume*

$$\lim_{x\to a}(f(x)) = b, lim_{x\to a}(g(x)) = c$$

then:

(i) $\lim_{x\to a}(f(x) + g(x)) = b + c$

(ii) $\lim_{x\to a}(\lambda f(x)) = \lambda b$

(iii) $\lim_{x\to a}(f(x) \cdot g(x)) = b \cdot c$

(iv) $\lim_{x\to a}|f(x)| = |b|$

Proof. of (iii)

$$\begin{aligned}
f(x) \cdot g(x) - b \cdot c &= f(x) \cdot g(x) - b \cdot g(x) + b \cdot g(x) - b \cdot c \\
&= g(x)\dot{(}f(x) - b) + b \cdot (g(x) - c) \\
|f(x) \cdot g(x) - b \cdot c| &= |g(x)\dot{(}f(x) - b) + b \cdot (g(x) - c)| \\
&\le |g(x)\dot{(}f(x) - b)| + |b \cdot (g(x) - c)|
\end{aligned}$$

Cauchy-Schwartz: $|x^1y^1 + \cdots + x^ny^n| \leq \sqrt{(x^1)^2 + \cdots + (x^n)^2} \cdot \sqrt{(y^1)^2 + \cdots + (y^n)^2}$

$$|f(x) \cdot g(x) - b \cdot c| \leq |g(x)\dot{(}f(x) - b)| + |b \cdot (g(x) - c)| \leq |g(x)| \cdot |f(x) - b| + |b| \cdot |g(x) - c|$$

Since $lim_{x \to a}(g(x)) = c$, g is a bounded neighbourhood of a, i.e:

$$\exists M \geq 0,\ \exists \delta > 0\ st,\ |g(x)| \leq M\ for\ |x - a| < \delta$$

□

Remark. *We have:*

(i) *$f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous iff: $f^i : \mathbb{R}^n \to \mathbb{R}$ is continuous for $i = 1, \ldots, m$*

(ii) *Polynomial functions in n-variables, $f(x^1, \ldots, x^n)$, are continuous*

(iii) *Rational functions, $R(x) = \frac{P(x)}{Q(x)}$, are continuous where defined, ie: $Q(x) \neq 0$ and P, Q are polynomials in n-variables.*

Theorem 1.2. *Linear transformations are continuous.*

Proof. $T : \mathbb{R}^n \to \mathbb{R}^m$ let $a \in \mathbb{R}^n$ to show: $\lim_{x \to a} T(a + h) = T(a)$, where $h = (h^1, \ldots, h^n)$

$$\begin{aligned} |T(a+h) - T(a)| = |T(h)| = |T(h^1e_1 + \cdots + h^ne_n)| &= |h^1T(e_1) + \cdots + h^nT(e_n)| \\ &\leq |h^1||T(e_1)| + \ldots |h^n||T(e_n)| \leq |h|(T(e_1) + \ldots T(e_n)) \end{aligned}$$

$$So: \quad |T(a+h) - T(a)| \leq M|h| \quad where \quad M = \sum_{i=1}^{n} |T(e_i)|$$

$$So\,given \quad \epsilon > 0, \quad choose \quad \delta = \frac{\epsilon}{M} \quad such\ that \quad |h| < \delta \Rightarrow |T(a+h) - T(a)| < \epsilon$$

□

Example 1.1. $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$, $\quad (x, y) = (0, 0)$ *assume* $\quad \lim_{(x,y) \to (0,0)} f(x, y) = L$

$$\forall \epsilon > 0, \quad \exists \delta > 0 \quad such\ that \quad 0 < |(x, y)| < \delta \Rightarrow |f(x, y) - L| < \epsilon$$

Plug $(x, 0)$ into f:

$$f(x, 0) \ = \frac{x^2 - 0}{x^2 - 0} = 1$$

Plug $(0, y)$ into f:

$$f(0, y) \ = \frac{0 - y^2}{0 + y^2} = -1$$

$$\begin{aligned} If\ |x| < \delta \quad |f(x, 0)| < \delta \Rightarrow |f(x, 0) - L| < \epsilon \quad ie \quad |1 - L| < \epsilon \\ If\ |y| < \delta \quad |f(0, y)| < \delta \Rightarrow |f(0, y) - L| < \epsilon \quad ie \quad |-1 - L| < \epsilon \\ \Rightarrow \epsilon = \frac{1}{2} \quad contradiction! \end{aligned}$$

Now consider $y = mx, m \in \mathbb{R}$

$$
\begin{aligned}
f(x, mx) &= \frac{x^2 - (mx)^2}{x^2 + (mx)^2} = \frac{1 - m^2}{1 + m^2} \\
\lim_{x \to 0}(\lim_{y \to 0} f(x, y)) &= \lim_{x \to 0} 1 = 1 \\
\lim_{y \to 0}(\lim_{x \to 0} f(x, y)) &= \lim_{y \to 0} -1 = -1
\end{aligned}
$$

However checking along straight lines is not enough to prove continuity.

Example 1.2.

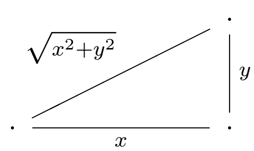
$$
f(x, y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}
$$

Show f is continuous at (0,0)

$$\forall \epsilon > 0, \quad \exists \delta > 0$$

$$
\left| \frac{xy}{\sqrt{x^2 + y^2}} \right| \leq \frac{|x| \cdot |y|}{\sqrt{x^2 + y^2}} \leq \frac{\sqrt{x^2 + y^2} \cdot \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}} = \sqrt{x^2 + y^2} = |(x, y)|
$$

Since:



Note. *if the total degree of the neumerator is higher than the denominator in a rational function. Then the limit should be 0.*

Theorem 1.3. *If f is continuous at a and g is continuous at $f(a)$ then $g \circ f$ is continuous at a.*

1.3 Partial Derivitives

Definition 1.3. *Let $f : \mathbb{R}^n \to \mathbb{R}$, $a \in \mathbb{R}$*

$$
Define: \quad D_i f(a) = \lim_{h \to 0} \frac{f(a^1, \ldots, a^{i-1}, a^i + h, a^{i+1}, \ldots, a^n)}{h}
$$

Example 1.3. *if $f : \mathbb{R}^n \to \mathbb{R}$*

$$
\left. \frac{df}{dx} \right|_{(a,b)} = D_1 f(a, b)
$$

$$
\left. \frac{df}{dy} \right|_{(a,b)} = D_2 f(a, b)
$$

$$
\textit{and in } \mathbb{R}^3 \textit{ we use } \frac{df}{dx}, \frac{df}{dy} \textit{ and } \frac{df}{dz} \textit{ etc.}
$$

Example 1.4.

$$f(x,y) = \begin{cases} \frac{x^2-y^2}{x^2+y^2} & \textit{if } (x,y) \neq (0,0) \\ 1 & \textit{if } (x,y) = (0,0) \end{cases}$$

$$D_1 f(0,0) = \left.\frac{df}{dx}\right|_{(0,0)} = \lim_{x\to 0} \frac{f(x,0)-f(0,0)}{x} = \lim_{x\to 0} \frac{\frac{x^2-0}{x^2-0} - 1}{x} = 0$$

$$D_2 f(0,0) = \left.\frac{df}{dy}\right|_{(0,0)} = \lim_{y\to 0} \frac{f(0,y)-f(0,0)}{y} = \lim_{y\to 0} \frac{\frac{0-y^2}{0+y^2} - 1}{y} = \frac{-2}{y} = \pm\infty$$

1.4 Total Derivitive

In 1 dimention we write the following for the derivitive of $f : \mathbb{R} \to \mathbb{R}$

$$f^{'}(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$$

we try to write it in higher dimentions $f : \mathbb{R}^n \to \mathbb{R}^m$ in this form

$$\begin{aligned} \lim_{h\to 0} \left[\frac{f(a+h)-f(a)}{h} - f^{'}(a) \right] &= \lim_{h\to 0} \left[\frac{f(a+h)-f(a)-h\cdot f^{'}(a)}{h} \right] \\ &= \lim_{h\to 0} \frac{|f(a+h)-f(a)-h\cdot f^{'}(a)|}{|h|} = 0 \end{aligned}$$

For $f : \mathbb{R}^n \to \mathbb{R}^m$ consider the tangent line at a: $y = f(a) + f^{'}(a)(x-a)$
call $x - a = h$ then we have: $y = f(a) + f^{'}(a)(h)$
this is an Affine transformation, not a linear map.
Look at the map:

$$\lambda : h \to hf^{'}(a), \quad h \in \mathbb{R}$$

This is a linear map.

$$\begin{aligned} \lambda(h_1+h_2) = (h_1+h_2)f^{'}(a) = h_1 f^{'}(a) + h_2 f^{'}(a) = \lambda(h_1) + \lambda(h_2) \\ \lambda(\alpha\cdot h) = (\alpha h) f^{'}(a) = \alpha(hf^{'}(a)) = \alpha\cdot\lambda(h) \\ \lim_{h\to 0} \frac{|f(a+h)-f(a)-\lambda(h)|}{|h|} = 0 \end{aligned}$$

Definition 1.4 (Total Derivitive)**.** *$f : \mathbb{R}^n \to \mathbb{R}^m$ or ($f : A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, A is open) is differentiable at a ($a \in A$) if we can rind a linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ st:*

$$\lim_{h\to 0} \frac{|f(a+h)-f(a)-\lambda(h)|}{|h|} = 0$$

The linear transformation λ is called the total derivitive of f at a and denoted Df(a) st

$$Df(a) = \lambda(h)$$

Example 1.5. *$f : \mathbb{R}^n \to \mathbb{R}^m$, $f(x) = k$, $k \in \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$ with the 0 linear transformation* $0 : f : \mathbb{R}^n \to \mathbb{R}^m$, $0(h) = 0$

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - 0(h)|}{|h|} = \lim_{h\to 0} \frac{|k - k - 0|}{|h|} = 0$$

Example 1.6. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformaton, it is differentiable at $a \in \mathbb{R}^n$ with linear transformation $Df(a) = f$*

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - f(h)|}{|h|} = \lim_{h\to 0} \frac{|f(a+h-a-h)|}{|h|} = 0$$

Theorem 1.4 (Uniqueness of Total Derivitive)**.** *If f is differentiable at a then there exists a unique linear transformation, $\lambda : \mathbb{R}^n \to \mathbb{R}^m$, such that*

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = 0$$

Proof. suppose $\mu : \mathbb{R}^n \to \mathbb{R}^m$ is another linear transformation such that:

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - \mu(h)|}{|h|} = 0$$

deduce that $\lambda = \mu \,\forall h \in \mathbb{R}^n$ *ie* $\lambda(h) = \mu(h)$

$$\begin{aligned}\frac{|\lambda(h) - \mu(h)|}{|h|} &= \frac{|\lambda(h) + f(a) - f(a+h) + f(a+h) - f(a) - \mu(h)|}{|h|}\\ &\leq \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} + \frac{|f(a+h) - f(a) - \mu(h)|}{|h|}\end{aligned}$$

Conclude that:

$$lim_{h\to 0} \frac{|\lambda(h) - \mu(h)|}{|h|} \leq 0 + 0 = 0 \quad (*)$$

Let h=0 $\lambda = 0 = \mu$ since λ, μ are linear. Now fix $h \in \mathbb{R}^n$, $h \neq 0$ and let $t \in \mathbb{R}$ such that $th \in \mathbb{R}^n$ then replace h with th in (*):

$$\begin{aligned}\lim_{t\to 0} \frac{|\lambda(th) - \mu(th)|}{|th|} &= \lim_{t\to 0} \frac{|t\lambda(h) - t\mu(h)|}{|t||h|}\\ &= \lim_{t\to 0} \frac{|t|}{|t|} \frac{|\lambda(h) - \mu(h)|}{|h|} = \frac{|\lambda(h) - \mu(h)|}{|h|} = 0\\ \lambda(h) = \mu(h) &\end{aligned}$$

□

Definition 1.5 (Jacobian Matrix)**.** *$f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$ and it is derivitive at a $Df(a) : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map. Then the matrix representation of $Df(a)$ is $f^{'}(a) \in \mathbb{M}_{mxn}$ and is called the Jacobian Matrix of f at a.*

Example 1.7. $f:\mathbb{R}^2\to\mathbb{R}^2, \quad f(x,y)=(x^2,x+5) \quad x,y\in\mathbb{R}$
Show that $Df(1,2)(h^1,h^2)=(4h^1+h^2,h^1)$*:*

$$\begin{aligned}
&f((1,2)+(h^1,h^2))-f(1,2)-Df(1,2)(h^1,h^2)\\
&=f(1+h^1,2+h^2)-f(1,2)-(4h^1+h^2,h^1)\\
&=((1+h^1)^2(2+h^2),(1+h^1+5))-(2,6)-(4h^1+h^2,h^1)\\
&=(2+h^2+2(h^1)^2+(h^1)^2h^2+2h^1h^2+4h^1-2-4h^1-h^2,6+h^1-6-h^1)
\end{aligned}$$

Take length:

$$|(2(h^1)^2+(h^1)^2h^2+2h^1h^2,0)|\leq 2|h|^2+|h|^2|h|+2|h||h|=4|h|^2+|h|^3$$

So:

$$\begin{aligned}
&\lim_{h\to 0}\frac{|f((1,2)+(h^1,h^2))-f(1,2)-Df(1,2)(h^1,h^2)|}{|h|}\\
&\leq \lim_{h\to 0}\frac{4|h|^2+|h|^3}{|h|}=\lim_{h\to 0}4|h|+|h|^2=0
\end{aligned}$$

Definition 1.6. $f^{'}(a)$ *is the matrix representation of* $Df(a)$

$$Df(a)(h)^t=\begin{pmatrix} y^1\\ y^2\\ \vdots\\ y^m\end{pmatrix}=f^{'}(a)\begin{pmatrix} h^1\\ h^2\\ \vdots\\ h^n\end{pmatrix}$$

$$f^{'}(a)=\begin{pmatrix} D_1f^1(a) & D_2f^1(a) & \cdots & D_nf^1(a)\\ D_1f^2(a) & D_2f^2(a) & \cdots & D_nf^2(a)\\ \vdots & \vdots & \ddots & \vdots\\ D_1f^m(a) & D_2f^m(a) & \cdots & D_nf^m(a)\end{pmatrix}$$

Example 1.8. *With this new information we can tackle example*

Remark. *Having directional derivitives in all directions* $u\neq 0$ *is not enough to guarantee* $df(a)$ *exists.*

Theorem 1.5. *If* f *is differentiable at* a *then* f *is continuous at* a*.*

Proof.

$$\begin{aligned}
\lim_{h\to 0}|f(a+h)-f(a)|&=\lim_{h\to 0}|f(a+h)-f(a)-Df(a)+Df(a)|\\
&\leq \lim_{h\to 0}\frac{|f(a+h)-f(a)-Df(a)(h)|}{|h|}\cdot|h|+\lim_{h\to 0}|Df(a)(h)|\\
&=0
\end{aligned}$$

since $Df(a)$ is a linear transformation $Df(a)$ is continuous so:

$$\lim_{h\to 0}|Df(a)(h)|=|Df(a)(0)|=0$$

. $\square$

1.5 The Chain Rule

Theorem 1.6 (Chain Rule)**.** *if $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a and $f : \mathbb{R}^m \to \mathbb{R}^k$ is differentiable at $f(a)$ then $g \circ f : \mathbb{R}^n \to \mathbb{R}^k$ is differentiable at a and*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$$

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "Df(a)"] \arrow[dr, "D(g\circ f)(a)"'] & \mathbb{R}^m \arrow[d, "Dg(f(a))"] \\
 & \mathbb{R}^k
\end{tikzcd}$$

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a), \quad \textit{where} \cdot \textit{represents matrix multiplication}$$

Proof. if $b = f(a)$ and we let $Df(a) = \lambda$ and $Dg(f(a)) = \mu$ then if we define:

$$\varphi(x) = f(x) - f(a) - \lambda(x - a) \tag{1}$$
$$\psi(y) = g(y) - g(b) - \mu(y - b) \tag{2}$$
$$\rho(x) = g \circ f(x) - g \circ f(a) - \mu \circ \lambda(x - a) \tag{3}$$

Then:

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - Df(a)(h)|}{|h|} = \lim_{x \to a} \frac{|\varphi(x)|}{|x - a|} = 0 \tag{4}$$
$$\lim_{h \to 0} \frac{|g(b+h) - g(b) - Dg(b)(h)|}{|h|} = \lim_{y \to b} \frac{|\psi(y)|}{|y - b|} = 0 \tag{5}$$

We must show:

$$\lim_{h \to 0} \frac{|g \circ f(x) - g \circ f(a) - \mu \circ \lambda(x - a)|}{|h|} = \lim_{x \to b} \frac{|\rho(x)|}{|x - b|} = 0$$

Now:

$$\begin{aligned}
\rho(x) &= g(f(x)) - g(b) - \mu(\lambda(x - a)) \\
&= g(f(x)) - g(b) - \mu(f(x) - f(a) - \varphi(x)) && \text{by (1)} \\
&= [g(f(x)) - g(b) - \mu(\lambda(f(x) - f(a)))] \\
&= \mu(\varphi(x)) = \psi(f(x)) + \mu(f(x)) && \text{by (2)}
\end{aligned}$$

Thus we must Prove

$$\lim_{x \to a} \frac{|\psi(f(x))|}{|x - a|} = 0 \tag{6}$$
$$\lim_{x \to a} \frac{|\mu\varphi(x)|}{|x - a|} = 0 \tag{7}$$

It follows from (5) that for some $\delta > 0$ we have

$$|\psi(f(x))| < \epsilon|f(x) - b| \quad \textit{if } |f(x) - b| < \delta$$

which is true if $|x-a| < \delta_1$ for a suitable δ_1. We also have that if T is a linear transformation then $\exists M \geq 0$ *such that* $|T(x)| < M|x|$. So then:

$$\begin{aligned} |\psi(f(x))| &< \epsilon|f(x) - b| \\ &= \epsilon|\varphi(x) + \lambda(x-a)| \\ &\leq \epsilon|\varphi(x)| + \epsilon M|x-a| \end{aligned}$$

So

$$\lim_{x\to a} \frac{|\psi(f(x))|}{|x-a|} \leq \lim_{x\to a} \frac{\epsilon|\varphi(x)|}{|x-a|} + \lim_{x\to a} \frac{\epsilon M|x-a|}{|x-a|} = \epsilon M \to 0$$

Also

$$\lim_{x\to a} \frac{|\mu\varphi(x)|}{|x-a|} \leq \lim_{x\to a} \frac{M|\varphi(x)|}{|x-a|} = 0$$

□

Theorem 1.7. *Define* $s : \mathbb{R}^2 \to \mathbb{R} \quad s(x, y) = x + y$ *then* s *is differentiable and* $Ds = s$

Proof. S is linear so

$$\begin{aligned} s((x, y) + (x', y')) &= s(x + x', y + y') = s(x, y) + s(x', y') \\ s(\lambda(x, y)) &= \lambda s(x, y) \end{aligned}$$

$$\lim_{h\to 0} \frac{|s(a+h) - s(a) - s(h)|}{|h|} = 0$$

□

Theorem 1.8. *Define* $p : \mathbb{R}^2 \to \mathbb{R}, \quad p(x, y) = xy$, *then* p *is differentiable and:*
$Dp(a, b) : \mathbb{R}^2 \to \mathbb{R}$ *is linear with* $Dp(a, b)(h, k) = ak + bh$ *and* $p' = (b, a)$

Proof. use of derivitive

$$\begin{aligned} p((a, b) + (h, k)) - p(a, b) - Dp(a, b)(h, k) &= p(a + h, b + k) - p(a, b) - (ak + bh) \\ &= (a + h)(b + k) - ab - (ak + bh) = hk \end{aligned}$$

$$\frac{|p((a, b) + (h, k)) - p(a, b) - Dp(a, b)(h, k)|}{|(h, k)|} = \frac{|hk|}{\sqrt{h^2 + k^2}} \leq \frac{\sqrt{h^2 + k^2}\sqrt{h^2 + k^2}}{\sqrt{h^2 + k^2}} = \sqrt{h^2 + k^2} \to 0$$

□

Remark. *To check some* $T : \mathbb{R}^n \to \mathbb{R}^m$ *is linear we listed two properties:*

$$\begin{aligned} T(x + y) &= T(x) + T(y) \\ T(\lambda x) &= \lambda T(x) \end{aligned}$$

we can instead just check:

$$T(\lambda x + y) = \lambda T(x) + T(y)$$

Corollary 1.1. $f, g : \mathbb{R}^n \to \mathbb{R}$ *differentiable at* $a \in \mathbb{R}^n$

(i) $D(f + g)(a) = Df(a) + Dg(a)$

(ii) Product rule: $D(f \cdot g)(a) = g(a) \cdot Df(a) + f(a) \cdot Dg(a)$

(iii) Quotient rule: if $g(a) \neq 0$, $D(\frac{f}{g})(a) = \frac{1}{g(a)^2} \cdot (g(a) \cdot Df(a) - f(a) \cdot Dg(a))$

Proof. For (i):
We can consder the function s from theorem

1.6 Mixed Derivitives

$f : \mathbb{R}^n \to \mathbb{R}, a \in \mathbb{R}$

$$D_i = lim_{h\to 0}\frac{f(a^1,\dots,a^{i-1},a^i+h,a^{i+1},\dots,a^n)-f(a)}{h}$$

if $D_if(x)$ exists for al a in some open set U then we get a function $U \xrightarrow{D_i} \mathbb{R},\ x \to D_if(x)$ then we can talk about partial derivitives of D_if eg $D_j(D_if(x)) = D_{ij}f(x)$
If $D_if(x)$ exists $\forall x \in U$ this is a function of x and we can consider $D_j(D_if(x)) = D_{ji}f(x)$
In general $i \neq j$ eg $f(x,y) = x^3y^5$:

$$\begin{array}{cc} D_1f(x,y) = 3x^2y^5 & D_2f(x,y) = 5x^3y^4 \\ D_{2,1}f(x,y) = 15x^2y^4 & D_{1,2}f(x,y) = 15x^2y^4 \end{array}$$

Theorem 1.9. *If $D_{i,j}$ and $D_{j,i}$ are continuous on an open set containing a then*

$$D_{i,j} = D_{j,i}$$

Proof. from homework 5:
First we repeat the well-known proof that, if $g : U \to \mathbb{R}$ is continuous and $g(p) > 0$, then there exists a neighborhood V of $p(p \in V \subset U,\ V\ open)$ with

$$q \in V \Rightarrow g(q) > 0$$

Take $\epsilon = g(p)$ in the defnition of continuity of g. There there exists a V open with $p \in V$ and

$$q \in V \Rightarrow |g(q) - g(p)| < g(p)$$

Since

$$g(p) - g(q) \leq |g(q) - g(p)| < g(p) \Rightarrow -g(q) < 0 \Leftrightarrow g(q) > 0$$

we get the result. The set V can be taken to contain a closed rectangle $[a,b]\times[c,d]$.
We apply the result to $g = D_{1,2}f - D_{2,1}f$. Assume (by contradiction) that $g(p)$ is not always 0. Then there exists a point p with $g(p) \neq 0$. We can assume that $g(p) > 0$, otherwise consider $-g$. The function g is given to be continuous. We have (using Fubini twice)

$$0 < \int_{[a,b]\times[c,d]} (D_{1,2}f(x,y) - D_{2,1}f(x,y))dA$$

$$= \int_a^b \left(\int_c^d D_{1,2}f(x,y)dy\right)dx - \int_a^b \left(\int_c^d D_{2,1}f(x,y)dx\right)dy$$

$$= \int_a^b (D_1f(x,d) - D_1f(x,c))\,dx - \int_c^d (D_2f(b,y) - D_2f(a,y))\,dy$$

$$= (f(b,d) - f(a,d) - f(b,c) + f(a,c)) - (f(b,d) - f(b,c) - f(a,d) + f(a,c)) = 0$$

using the fundamental theorem of calculus 6 times. This is a contradiction, so the mixed partial derivatives are equal on the rectangle. ∎

Theorem 1.10. *$A \subset \mathbb{R}$ If the max or min of $f : A \to \mathbb{R}$ occur at a point a in the interior of A and $D_if(x)$ exists then $Df(a) = 0$*

Proof. Consider $h(x) = f(a^1, \dots, a^{i-1}, x^i, a^{i+1}, \dots a^n)$ x in an open interval arround a^i. Since f has a max or min at a, h has a max or min at a^i

$$\frac{dh}{dx}(a^i) = D_i f(a)$$

By analysis 2:

$$\frac{dh}{dx}(a^i) = 0 \Rightarrow Df(a) = 0$$

□

Note. *The converse of Theorem* **??** *is not true, even in one dimension.*

1.7 Jacobian

For $f : \mathbb{R}^n \to \mathbb{R}^m$ with total derivitive $Df(a) : \mathbb{R}^n \to \mathbb{R}^m$ a linear map. Then the Jacobian $f^{'}(a) \in \mathbb{M}_{mxn}$ is the unique representation of $Df(a)$ in the standard basis.

Theorem 1.11. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a then $D_i f^j(a)$ exists $\forall i = 1, \dots, n$ $\forall j = 1, \dots, m$ and the jacobian matrix is*

$$f^{'}(a) = (D_i f^j(a))^{j=1,\dots,m}_{j=1,\dots,n}$$

$$f^{'}(a) = \begin{pmatrix} D_1 f^1(a) & D_2 f^1(a) & \cdots & D_n f^1(a) \\ D_1 f^2(a) & D_2 f^2(a) & \cdots & D_n f^2(a) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f^m(a) & D_2 f^m(a) & \cdots & D_n f^m(a) \end{pmatrix}$$

Where $f(x) = (f^1(x), \dots, f^m(x))$, $f^i : \mathbb{R}^n \to \mathbb{R}$

Proof. Case $m = 1$

$$\begin{tikzcd} \mathbb{R} \arrow[r, "h"] \arrow[dr, "f \circ h"'] & \mathbb{R}^n \arrow[d, "f"] \\ & \mathbb{R} \end{tikzcd}$$

$$h(t) = (a^1, \dots, a^{i-1}, t, a^{i+1}, \dots, a^n) \qquad \left.\frac{d(f \circ h)}{dt}\right|_{t=a^i} = D_i f(a)$$

$$\lim_{t \to a^i} \frac{(f \circ h)(t) - (f \circ h)(a^i)}{t - a^i} = \lim_{t \to a^i} \frac{f(a^1, \dots, a^{i-1}, t, a^{i+1}, \dots, a^n) - f(a^1, \dots, a^n)}{t - a^i}$$

h is differentiable because its components are differentiable ie component h^i is either constant a^j where $j \neq i$ or t when $j = i$

$$\begin{aligned} Dh(t) &= (Dh^1(t), \dots, Dh^n(t)) \\ &= (0, \dots, 1, \dots, 0) \end{aligned}$$

$$h^{'}(a^i) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}_{(m\times 1)}$$

Case $m > 1$
$f : \mathbb{R}^n \to \mathbb{R}^m$

$$f(x) = (f^1(x), \dots f^m(x))$$
$$Df(a) = (Df^1(a), \dots Df^a)$$
$$f^{'}(a) = \begin{pmatrix} (f^1)^{'}(a) \\ \vdots \\ (f^m)^{'}(a) \end{pmatrix}_{(m\times n)}$$
$$f^{'}(a) = \begin{pmatrix} D_1f^1(a) & D_2f^1(a) & \cdots & D_nf^1(a) \\ D_1f^2(a) & D_2f^2(a) & \cdots & D_nf^2(a) \\ \vdots & \vdots & \ddots & \vdots \\ D_1f^m(a) & D_2f^m(a) & \cdots & D_nf^m(a) \end{pmatrix}$$

□

Remark. *Abuse of notation since if $f : \mathbb{R} \to \mathbb{R}$*

$$\textit{this is a number} \to \frac{dg(t_0)}{dt}| = g'(t_0) \leftarrow \textit{this is the } 1 \times 1 \textit{ jacobian matrix}$$

Example 1.9.

$$G(x, y) = \begin{cases} \frac{x^2y}{x^4+y^2} & \textit{if } (x, y) \neq (0, 0) \\ 0 & \textit{if } (x, y) = (0, 0) \end{cases}$$

Fix a vector $u \in \mathbb{R}^2$, $u = (u^1, u^2) \neq (0, 0)$, $u^2 \neq 0$ then the directional derivitive D_u with $h \in \mathbb{R}$ is:

$$\begin{aligned} D_uG(0, 0) &= \lim_{h\to 0} \frac{G((0, 0) + hu) - G(0, 0)}{h} = \lim_{h\to 0} \frac{G(hu^1, hu^2) - 0}{h} \\ &= \lim_{h\to 0} \frac{(hu^1)^2(hu^2)}{(hu^1)^4 + (hu^2)^2} \cdot \frac{1}{h} = \lim_{h\to 0} \frac{h^3(u^1)^2u^2}{h(h^4(u^1)^4 + h^2(u^2)^2)} \\ &= \lim_{h\to 0} \frac{(u^1)^2u^2}{h^2(u^1)^4 + (u^2)^2)} = \frac{(u^1)^2}{u^2} \end{aligned}$$

$u^2 = 0$

$$D_uG(0, 0) = \lim_{h\to 0} \frac{G(hu^1, h\cdot 0)}{h} = \lim_{h\to 0} \frac{\left(\frac{(hu^1)^2 0}{(hu^1)^4+0^2}\right)}{h} = 0$$

Theorem 1.12. *$f:\mathbb{R}^n \to \mathbb{R}^m$ if $D_jf^i(x)$ exist $\forall x \in u, U$ open, $a \in U, \forall i = 1,\ldots,m$, and $j = 1,\ldots,n$ and if $D_jf^i(x)$ continuous at a ie*

$$\lim_{x\to a}(D_jf^i(x)) = D_jf^i(a)$$

then $Df(a)$ exists and f is differentiable at a

Proof. As in the proof of theorem **??** It suffices to consider the case $m = 1$, so that $f:\mathbb{R}^n \to \mathbb{R}$. Then

$$\begin{aligned} f(a+h) - f(a) = {} & f(a^1+h^1, a^2,\ldots,a^n) && - f(a^1,\ldots,a^n) \\ & + f(a^1+h^1, a^2+h^2, a^3,\ldots,a^n) && - f(a^1+h^1, a^2,\ldots,a^n) \\ & + \ldots && - \ldots \\ & + f(a^1+h^1,\ldots,a^n+h^n) && - f(a^1+h^1,\ldots,a^{n-1}+h^{n-1},a^n) \end{aligned}$$

Recal from theorem **??** that D_1f is the derivitive of the function h defined by $h(x) = (x, a^2,\ldots,a^n)$. Applying the mean-value theorem to h we obtain

$$f(a^1+h^1, a^2,\ldots,a^n) - f(a^1,\ldots,a^n) = h^1 \cdot D_1f(b_1, a^2,\ldots,a^n)$$

for some b_1 between a^1 and a^1+h^1. Similarly the *ith* term in the sum equals

$$h^i \cdot D_if(a^1+h^1,\ldots,a^{i-1}+h^{i-1}, b_i, a^{i+1},\ldots,a^n) = h^iD_if(c_i) \quad \text{for some } c_i$$

Then

$$\begin{aligned} \lim_{h\to 0}\frac{|f(a+h) - f(a) - \sum_{i=1}^n D_if(a)\cdot h^i|}{|h|} &= \lim_{h\to 0}\frac{|\sum_{i=1}^n[D_if(c_i) - D_if(a)]\cdot h^i|}{|h|} \\ &\le \lim_{h\to 0}|\sum_{i=1}^n[D_if(c_i) - D_if(a)]|\cdot\frac{|h^i|}{|h|} \\ &\le \lim_{h\to 0}|\sum_{i=1}^n[D_if(c_i) - D_if(a)]| \\ &= 0 \end{aligned}$$

Since D_if is continuous at a and as $h \to 0, c^i \to a^i$. □

Definition 1.7. *If $f:\mathbb{R}^n \to \mathbb{R}^m$ has partial derivitives $D_jf^i \quad \forall x \in U$, U open, $a \in U$ and D_jf^i is continuous at a then we say f is continuously differentiable at a.*

Example 1.10. *$f:\mathbb{R}^2 \to \mathbb{R}$, with $x:\mathbb{R}\to\mathbb{R}, y:\mathbb{R}\to\mathbb{R}$.*

$$\textit{Define} \quad g:\mathbb{R}\to\mathbb{R} \quad g(t) = f(x(t), y(t))$$

$$\begin{aligned} \frac{dg(t_0)}{dt} &= (g'(t_0)) = f'(x(t_0), y(t_0)) \cdot \begin{pmatrix} x'(t_0) \\ y'(t_0) \end{pmatrix} \\ &= \frac{df}{dx}(x(t_0), y(t_0)) + \frac{df}{dy}(x(t_0), y(t_0)) \\ &= \frac{df}{dx}(x(t_0), y(t_0)) \cdot \frac{dx}{dt}(t_0) + \frac{df}{dy}(x(t_0), y(t_0)) \cdot \frac{dy}{dt}(t_0) \\ &= \frac{df}{dx}\cdot\frac{dx}{dt} + \frac{df}{dy}\cdot\frac{dy}{dt} \end{aligned}$$

1.8 20/11/11 some deep graphs in romans notes, may have a crack producing them later on

1.9 Inverse Function Theorem

Lemma 1.13. *Let $A \subset \mathbb{R}^n$ be a rectangle with interior A^0 and let $g : A \to \mathbb{R}^n$ be continuously differentiable. If there exist a constant $M > 0$ such that*

$$|D_j g^i(x)| \leq M, \quad x \in A^0, \quad i, j = 1, \ldots, n.$$

then

$$|g(x) - g(y)| \leq n^2 M |x - y|, \quad x, y \in A.$$

Proof. Fix $i = 1, \ldots, n$. Then

$$\begin{aligned} g^i(y) - g^i(x) &= g^i(y^1, y^2, \ldots, y^n) - g^i(x^1, x^2, \ldots, x^n) \\ &= g^i(y^1, y^2, \ldots, y^n) - g^i(x^1, y^2, \ldots, y^n) + g^i(x^1, y^2, \ldots, y^n) - g^i(x^1, x^2, \ldots, y^n) \\ &\quad + g^i(x^1, x^2, \ldots, y^n) - \cdots + g^i(x^1, x^2, \ldots, y^n) - g^i(x^1, x^2, \ldots, x^n) \\ &= \sum_{j=1}^n (g^i(x^1, x^2, \ldots, x^{j-1}, y^j, \ldots, y^n) - g^i(x^1, x^2, \ldots, x^{j-1}, x^j, y^{j+1}, \ldots, y^n) \\ &= \sum_{j=1}^n (y^j - x^j) D_j g^i(z^i_j) \end{aligned}$$

where z^i_j is between y^j and x^j, and we used the mean-value theorem in the interval between y_j and x_j and in the j variable. Using the triangle inequality and $|z^j| \leq |z|$, we get

$$|g^i(y) - g^i(x)| \leq \sum_{j=1}^n |y^i - x^i| M \leq \sum_{j=1}^n |y - x| M = nM|y - x|.$$

Since $|z| \leq \sum_i |z^i|$, finally we get

$$|g(x) - g(y)| \leq \sum_{i=1}^n |g^i(y) - g^i(x)| \leq \sum_{i=1}^n nM|y - x| = n^2 M |y - x|.$$

□

Remark. *It is clear that the dimension of the target space enters only in the last line of the calculation. If $g : \mathbb{R}^n \to \mathbb{R}^m$, then we get as upper bound $nmM|x - y|$. The inequality is actually not optimal: one can use the Cauchy-Schwarz inequality twice to get a bound $n^{1/2} m^{1/2} M |x - y|$ for $g : \mathbb{R}^n \to \mathbb{R}^m$.*

Theorem 1.14 (Inverse Function Theorem)**.** *Theorem Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable on an open set containing a and assume $det f'(a) \neq 0$. Then there exists an open set V containing a and an open set W containing $f(a)$ such that $f : V \to W$ is bijective with $f^{-1} : W \to V$ continuously differentiable and which satisfies:*

$$(f^{-1})'(y) = [f'(f^{-1}(y))]^{-1}, \quad y \in W.$$

Proof. Step 1:
We reduce proving the theorem to the case where actually $f^{'}(a) = I_{nn}$. Call $\lambda = Df(a)$. This is a linear transformation with nonsingular matrix representation $f^{'}(a)$, as $\det f^{'}(a) \neq 0$. Therefore, λ is invertible. The inverse λ^{-1} is also a linear transformation, so $D(\lambda^{-1})(y) = \lambda^{-1}$ for $y \in \mathbb{R}^n$. Both λ and its inverse are continuous as linear transformations. Consider the function $h = \lambda^{-1} \circ f$ defined on an open set comtaining a.
Then:

$$Dh(a) = D\lambda^{-1}(f(a)) \circ Df(a) = \lambda^{-1} \circ Df(a) = \lambda^{-1} \circ \lambda = Id,$$

by using the chain rule. Here Id is the identity transformation. This gives $h^{'}(a) = I_{n\times n}$, which has determinant $1 \neq 0$. Let A be the matrix representation of λ^{-1}. (which gives that A^{-1} is the matrix representation of $\lambda = D\lambda$). This is an $n \times n$ matrix with constant entries, i.e. not depending on y. Moreover, h is continuously differentiable, as

$$(D_j h^i(x)) = h^{'}(x) = [\lambda^{-1} \circ Df(x)] = A \cdot f^{'}(x) = A(D_j f^i(x)),$$

with entries depending continuously on x. Therefore, h satisfies the conditions of the inverse function theorem. Suppose that we can prove the conclusion of it for h, i.e. that there exists an open set V containing a and $\tilde{W}$ open containing $h(a) = \lambda^{-1}(f(a))$ such that $h : V \to \tilde{W}$ is bijective with continuously differentiable inverse h^{-1}. Even more, assume that we have prove the formula for the derivative of the inverse of h:

$$(h^{-1})'(z) = [h'(h^{-1}(z))]^{-1}.$$

Define $W = \lambda(\tilde{W}) = (\lambda^{-1})^{-1}(\tilde{W})$. This is the inverse image of $\tilde{W}$ by λ^{-1}, which is continuous, so it is an open set. Since λ is bijective, $f = \lambda \circ h$ is bijective on V with image $\lambda(\tilde{W}) = W$. Moreover,

$$f^{-1} = h^{-1} \circ \lambda^{-1},$$

which is continuously differentiable as the composition of two such maps. By the chain rule for Jacobian

$$(f^{-1})'(y) = (h^{-1})'(\lambda^{-1}(y)) \cdot (\lambda^{-1})'(y) = [h'(h^{-1}(\lambda^{-1}(y)))]^{-1}A = [h'((\lambda \circ h)^{-1}(y))]^{-1}A = [h'(f^{-1}(y))]^{-1}A$$

$$= [A^{-1}h'(f^{-1}(y))]^{-1} = [\lambda' h'(f^{-1}(y))]^{-1} = [(\lambda \circ h)'(f^{-1}(y))]^{-1} = [f'(f^{-1}(y))]^{-1}.$$

All these imply that it is enough to work with $h = \lambda^{-1} \circ f$. The main property we will use is that $h'(a) = I_{n\times n}$. For simplicity in our notation we call this function f so we can assume that

$$f'(a) = I_{n\times n}.$$

This also means that $\lambda = Df(a) = Id$.
Step 2: The function f cannot take the value $f(a)$ arbitrarily close to a. Suppose that there is a sequence $h_n \in \mathbb{R}^n$ such that $h_n \to 0$ and $f(a + h_n) = f(a)$. We plug the sequence into the definition of the derivative at a and use that $Df(a) = Id$ to get

$$0 = \lim_{h_n \to 0} \frac{|f(a + h_n) - f(a) - Df(a)(h_n)|}{|h_n|} = \lim_{h_n \to 0} \frac{|-h_n|}{|h_n|} = 1$$

So this is a contradiction. Therefore, we can find a closed rectangle U containing a such that

$$f(x) \neq f(a), \quad \forall x \in U \backslash \{a\}.$$

Step 3: The determinant is a polynomial expression in the entries of a matrix. If the matrix entries depend continuously on x, the same is true for the determinant of the matrix. So $\det f'(x)$ is a continuous function on an open set containing a. Since $\det f'(x) \neq 0$, by the inertia principle, there exists a small enough (rectangular) neighbourhood of a, which we call U again, such that

$$\det f'(x) \neq 0, \quad x \in U \tag{1}$$

Moreover the partial derivatives $D_j f^i(x)$ are continuous and $D_j f^i(a) = \delta_{ij}$, as $Df(a) = Id$. So, for x close enough to a we have

$$|D_j f^i(x) - \delta_{ij}| < \frac{1}{2n^2}, \quad i, j = 1, \ldots, n, \quad x \in U \tag{2}$$

We assumed again that the neighbourhood is U

Step 4: Constructing a contraction map and showing that f is injective in appropriate small meighbourhood. Now we define the function

$$g(x) = f(x) - x$$

and apply the Lemma to this function for the closed rectangle U. We notice that $D_j g^i(x) = D_j f^i(x) - \delta_{ij}$, as we know the partial derivatives of the identity function x. We deduce that

$$|g(x_1) - g(x_2)| \leq n^2 \frac{1}{2n^2} |x_1 - x_2| = \frac{1}{2}|x_1 - x_2| \tag{3}$$

The choice of the neighbourhood in (2) so that the constant $1/(2n^2)$ appears on the right is motivated with the desire to get g as a contraction map (with constant $1/2$) as we see in (3). Now the triangle inequality in the form $|a| - |b| \leq |a - b|$ gives

$$|x_1 - x_2| - |f(x_1) - f(x_2)| \leq |(x_1 - x_2) - (f(x_1) - f(x_2))| = |-g(x_1) + g(x_2)| < \frac{1}{2}|x_1 - x_2|$$

$$\Rightarrow |x_1 - x_2| - \frac{1}{2}|x_1 - x_2| < |f(x_1) - f(x_2)| \Rightarrow \frac{1}{2}|x_1 - x_2| < |f(x_1) - f(x_2)| \tag{4}$$

Here x_1, x_2 are in U. We immediately see that on U the function f is injective:

$$f(x_1) = f(x_2) \Rightarrow |x_1 - x_2| = 0 \Rightarrow x_1 = x_2.$$

We still have not determined the neighbourhoods W of $f(a)$ and V of a.

Step 5: Determination of the minimum distance of $f(a)$ to the image of the boundary of U and definition of W.

We have assumed that on the closed rectangle U we have $f(x) \neq f(a)$ for $x \neq a$. This is definitely true on the boundary of U, denoted ∂U, which is a closed and bounded set, i.e. compact. The function $m(x) = |f(x) - f(a)|$ is continuous on a neighbourhood of ∂U and nonzero on it. It achieves a minimum value on ∂U (an advanced argument from Real Analysis is that the image of a compact set is compact, so that $m(\partial U)$ is compact, which means closed and bounded. Such a set has a maximum and minimum). The minimum value cannot be zero, say

$$\min_{x \in \partial U} m(x) = \min_{x \in \partial U} |f(x) - f(a)| > 0.$$

Now define

$$W = \{y \in \mathbb{R}^n, |y - f(a)| < \delta/2\}.$$

Step 6: Comparison of $|y - f(x)|$ with $|y - f(a)|$ for $x \in \partial U$, and $y \in W$. We have

$$|f(x)-f(a)| \geq \delta, \quad |y-f(a)| \leq \delta/2 \Rightarrow -|y-f(x)|+\delta \leq -|y-f(x)|+|f(x)-f(a)| \leq |y-f(a)| < \delta/2$$

$$\Rightarrow \delta/2 = \delta - \delta/2 < |y - f(x)| \Rightarrow |y - f(a)| < \delta/2 < |y - f(x)|.$$

Step 7: Show that for $y_0 \in W$ there exists a unique $x_0 \in U^0$ such that $f(x_0) = y_0$. The uniqueness is obvious from the fact that f is injective on U. The construction of such an x_0 is tricky. We define another function on U by

$$g(x) = |f(x) - y_0|^2 = \sum_{i=1}^{n} (f^i(x) - y_0^i)^2.$$

This function in continuously differentiable, as it is a sum of the squares of the components. On the compact set U the function g achieves its minimum, say at x_0, i.e. $g(x_0) \leq g(x)$ for $x \in U$. We claim that x_0 is the desired point with $f(x_0) = y_0$. First we see that x_0 is in the interior of the set U. On the boundary of U the function $g(x)$ has values $> \delta/2$, by Step 6, while $g(a) < \delta/2$. So the minimum is not achieved on the boundary of U. Therefore, it is achieved in an interior point. This point has to be a critical point of g, i.e. $D_j g(x_0) = 0,\ j = 1, \ldots, n$. We calculate them to be

$$2 \sum_{i=1}^{n} (f^i(x_0) - y_0^i) D_j f^i(x_0) = 0, \quad j = 1, \ldots, n.$$

This is a homogeneous system of linear equations with unknowns $f^i(x_0) - y_0^i$ and coefficients $D_j f^i(x_0)$. The determinant of the coefficients of the system is nonzero, as $x_0 \in U$. The system has a unique solution, and this solution is the zero vector, i.e

$$0 = f^i(x_0) - y_0^i, \quad i = 1, \ldots, n \Rightarrow f(x_0) = y_0.$$

Step 8: We define V and Show that $f : V \to W$ is bijective and continuous. We define $V = U^0 \cap f^{-1}(W)$. Clearly $f : V \to W$ is bijective. Moreover, V is open as the intersection of the open set U^0 and the open set $f^{-1}(W)$, which is open as the inverse image of an open set W by the continuous function f. We now rewrite (4) as

$$|x_1 - x_2| < 2|f(x_1) - f(x_2)| \Leftrightarrow |f^{-1}(y_1) - f^{-1}(y_2)| < 2|y_1 - y_2| \tag{5}$$

with $y_1 = f(x_1)$ and $y_2 = f(x_2)$, $y_i \in W$. This shows that f^{-1} is a Lipschitz function with constant 2, so that it is continuous. Alternatively, choose $\delta = \epsilon/2$ in the definition of continuity.

Step 9: Show that f^{-1} is differentiable. Let $\mu = Df(x_1)$. Since $f^{-1} \circ f = Id$, the chain rule gives the only possible choice for $Df^{-1}(y_1) = \mu^{-1}$. Here $f(x_1) = y_1$ and later $f(x) = y$. By the definition of the derivative we have

$$f(x) - f(x_1) = \mu(x - x_1) + \phi(x - x_1), \quad \lim_{x \to x_1} \frac{|\phi(x - x_1)|}{|x - x_1|} = 0.$$

We apply to the equation the linear transformation μ^{-1} to get

$$\mu^{-1}(y - y_1) = x - x_1 + \mu^{-1}(\phi(x - x_1)) \Rightarrow x - x_1 - \mu^{-1}(y - y_1) = \mu^{-1}(\phi(x - x_1))$$

$$\Rightarrow f^{-1}(y) - f^{-1}(y_1) - \mu^{-1}(y - y_1) = -\mu^{-1}(\phi(x - x_1)).$$

By the definition of the derivative of f^{-1} at y_1 we need to show that

$$\lim_{y \to Y_1} \frac{|-\mu^{-1}(\phi(x - x_1))|}{|y - y_1|} = 0 \tag{6}$$

Since μ^{-1} is a linear transformation, we have seen that it is a bounded linear operator, i.e. there exists a constant $\tilde{M}$ with

$$|\mu^{-1}(y)| \leq \tilde{M}|y|, \quad \forall y \in \mathbb{R}^n.$$

Since

$$\frac{|-\mu^{-1}(\phi(x - x_1))|}{|y - y_1|} \leq \frac{\tilde{M}|\phi(x - x_1)|}{|y - y_1|}$$

by the sandwich theorem it is enough to prove that

$$\lim_{y \to Y_1} \frac{|\phi(x - x_1)|}{|y - y_1|} = 0$$

We have

$$\frac{|\phi(x - x_1)|}{|y - y_1|} = \frac{|\phi(x - x_1)|}{|x - x_1|} \frac{|x - x_1|}{|y - y_1|} \leq \frac{|\phi(x - x_1)|}{|x - x_1|} \cdot 2,$$

by (5). Moreover, $y \to y_1$ iff $x \to x_1$ as f is continuous at x_1 and f^{-1} is continuous at y_1.We know that

$$\lim_{x \to x_1} \frac{|(\phi(x - x_1))|}{|x - x_1|} = 0$$

This suffices to prove (6)

Step 10: The partial derivatives $D_j(f^{-1})^i(y)$ are continuous. We know that the Jacobian of $f^{-1}(y)$ is

$$(f^{-1})'(y) = (D_j(f^{-1})^i(y)) = [f'(f^{-1}(y))]^{-1} = (D_j f^i(x))^{-1}.$$

The inverse of the matrix $(D_j f^i(x))$ can be calculated as a quotient of two $n \times n$ determinants with entries among $D_j f^i(x)$. The denominator is the determinant of the Jacobian at x, which is nonzero for $x \in U$. The whole expression depends continuously on $x \in V$. As f^{-1} is continuous, the inverse matrix depends continuously on $y \in W$. The individual entries are the partial derivatives of f^{-1}. □

1.10 Implicit Function Theorem

Example 1.11.

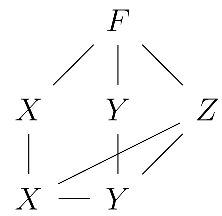
$$x^2 + y^2 = 1, \quad y = g(x)$$

$$2x + 2y\frac{dy}{dx} = 0, \quad \frac{dy}{dx} = \frac{dg}{dx} = \frac{-x}{y}, \quad y \neq 0$$

Example 1.12.

$$y^2 + xz + z^2 - e^z - 4 = 0 \quad \textit{(impossible to solve for z)}$$

$$\textit{set } F(x,y,z) = y^2 + xz + z^2 - e^z - 4, \quad F(x,y,g(x,y)) = 0$$



Differentiate in x:

$$\begin{aligned}
\frac{d}{dx}F(x,y,g(x,y)) &= \frac{dF}{dx}\frac{dx}{dx}+\frac{dF}{dy}\frac{dy}{dx}+\frac{dF}{dz}\frac{dz}{dx}\\
&= \frac{dF}{dx}+\frac{dF}{dz}\frac{dg}{dx}=0\\
\frac{dg}{dx} &= -\frac{\frac{dF}{dx}}{\frac{dF}{dz}}=-\frac{z}{x^2+2z-e^z}\\
\frac{dF}{dy}=0 \overset{chain}{\underset{rule}{\Rightarrow}} \frac{dF}{dy}+\frac{dF}{dz}\frac{dg}{dy} \Rightarrow \frac{dg}{dy} &= -\frac{\frac{dF}{dy}}{\frac{dF}{dz}}=-\frac{2y}{x^2+2z-e^z}
\end{aligned}$$

the point $(0,e,2)$ satisfies $F(x,y,z)=0$

$$e^2+0\cdot 2+2^2-e^2-4=0$$

$$\begin{aligned}
\left.\frac{dg}{dx}\right|_{(0,e)} &= -\frac{z}{x^2+2z-e^z}=-\frac{2}{0+2\cdot 2-e^2}\\
\left.\frac{dg}{dy}\right|_{(0,e)} &= -\frac{2y}{x^2+2z-e^z}=-\frac{2e}{0+2\cdot 2-e^2}
\end{aligned}$$

valid for $\frac{dF}{dz}\neq 0$

General situation: m equations with m unknowns $y^1,\dots,y^m$

$$\begin{array}{ll}
f^1(x^1,\dots,x^n,y^1,\dots,y^m)=0 & \text{depends on n parameters: } x^1,\dots,x^n\\
f^2(x^1,\dots,x^n,y^1,\dots,y^m)=0 & \text{Try to solve for: } y^1,\dots,y^m\\
\vdots \qquad \vdots \qquad \vdots & \\
f^m(x^1,\dots,x^n,y^1,\dots,y^m)=0 &
\end{array}$$

$$x=(x^1,\dots,x^n),\qquad y=(y^1,\dots,y^m)$$

So we have:

$$\begin{aligned}
f^1(x,y)&=0\\
f^2(x,y)&=0\\
&\vdots\\
f^m(x,y)&=0
\end{aligned}$$

Define $f(x,y) = (f^1(x,y),\dots,f^m(x,y)) = \underset{\substack{\sqcup \\ vector}}{0} = \underbrace{(0,\dots,0)}_{m}$

Let $a \in \mathbb{R}^n,\ b \in \mathbb{R}^m$ such that $f(a,b,) = 0$ when we can find for each $(x^1,\dots,x^n)$ near $a = (a^1,\dots,a^n)$ a unique $y = (y^1,\dots,y^m)$ near $b = (b^1,\dots,b^m)$ such that: $f(x,y) = 0$, $f(x^1,\dots,x^n,y^1,\dots,y^m) = 0$

Theorem 1.15 (Implicit Function Theorem)**.** *$f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ continuously differentiable on an open set containing (a,b), $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. moreover $f(a,b) = 0$
consider the matrix*

$$M = (D_{j+n}f^i(a,b))^{i=1,\dots,m}_{j=1,;m}$$

assume $detM \neq 0$. Then there exist two open sets $A \subset \mathbb{R}^n$, $b \subset \mathbb{R}^m$, $a \in A$, $b \in B$. such that $\forall x \in A, \exists$ unique $g(x) \in B$ such that $f(x,g(x)) = 0$ Moreover $g : A \to B$ is differentiable.

Proof. Increase the dimension of the target. Define $F : \underbrace{U}_{in\ \mathbb{R}^n \times \mathbb{R}^m} \to \mathbb{R}^n \times \mathbb{R}^m$

$$F(x^1,\dots,x^n,y^1,\dots,y^m) = (x^1,\dots,x^n,f^1(x,y),\dots,f^m(x,y))$$

$$F(x,y) = (x,f(x,y))$$

F is continuously differentiable because $x^1,\dots,x^n$ are continuously differentiable and $f^1(x,y),\dots,f^m(x,y)$ are continuously differentiable (because $f(x,y)$ is continuously differentiable)

$$F(a,b) = (a,f(a,b)) = (a,0)$$

$$F'(a,b) = \left(\begin{array}{cccc|ccc}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\
\hline
\frac{df^1}{dx^1} & \frac{df^1}{dx^2} & \cdots\ \frac{df^1}{dx^n} & \frac{df^1}{dy^1} & \cdots & \frac{df^1}{dy^m} & \\
\vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\
\frac{df^m}{dx^1} & \frac{df^m}{dx^2} & \cdots\ \frac{df^m}{dx^n} & \frac{df^m}{dy^1} & \cdots & \frac{df^m}{dy^m} &
\end{array}\right)$$

$$F'(a,b) = \left(\begin{array}{c|c}
I_{n\times n} & 0_{n\times m} \\
\hline
\text{* Some } m\times n & M_{m\times m} \\
\text{matrix} &
\end{array}\right)$$

$\det M \neq 0$ (reducing from top left entry).
By the inverse funct theorem, $\exists$ an open set W containing $F(a,b) = (a,0)$ and an open set containing $(a,0)$ which I can take to be a rectangle $A \times B$, $a \in A$, $b \in B$, A open in $\mathbb{R}^n$, B open in $\mathbb{R}^m$.

$$F : A \times B \to W \text{ is bijective}$$

$$\exists h = F^{-1} : W \to A \times B \text{ such that } F \cdot h = id$$

h is continuously differentiable.

$$\boxed{\begin{gathered} F(x^1,\dots,x^n,y^1,\dots,y^m) = (x^1,\dots,x^n,f^1(x,y),\dots,f^m(x,y)) \\ F(x,y) = (x,f(x,y)) \\ F \text{ is continuously differentiable because } x^1,\dots x^n \text{ are continuously differentiable} \end{gathered}}$$

h must have the form: $h(x,y) = (x, k(x,y))$ for some function $k : W \to B,\ B \subset \mathbb{R}^m,\ k$ continuously differentiable.

$$F(h(x,y) = (x, f(x, k(x,y))) = (x,y)$$

$$f(x, k(x,y)) = y$$

Set $y = 0$

$$f(x, k(x,0)) = 0$$

The solution is $g(x) = k(x,0)$ (solution to $f(x,y) = 0$). $\square$

Theorem 1.16. *Let $g : \mathbb{R}^n \to \mathbb{R}^p$ be a continuously differentiable function in an open set containing a and assume that $p \leq n$. If $g(a) = 0$ and the rank of the $p \times n$ matrix*

$$(D_j g^i(a))_{i=1,\ldots,p\ j=1,\ldots,n}$$

be equal to p. Then there exists an open set $A \subset \mathbb{R}^n$ and a differentiable function $h : A \to \mathbb{R}^n$ which is bijective onto an open set V and h^{-1} is differentiable and

$$(g \circ h)(x^1, x^2, \ldots, x^n) = (x^{n-p+1}, x^{n-p+2}, \ldots, x^n)$$

Proof. We can consider the function g as $g : \mathbb{R}^{n-p} \times \mathbb{R}^P \to \mathbb{R}^p$. The 'easy' case is as follows: If the $p \times n$ matrix above is such that the last p columns give a matrix M with $det(M) \neq 0$, then we are exactly in the situation of the Implicit Function Theorem as worked out above. The notation has only slightly changed: $x^{n-p+1} = y^1,\ x^{n-p+2} = y^2, \ldots, x^n = y^p,\ p = m,\ g = f$. We have found h with $h(x,y) = (x, k(x,y))$ and

$$(f \circ h)(x,y) = f(h(x,y)) = f(x, k(x,y)) = y,$$

and in our notation

$$(g \circ h)(x^1, x^2, \ldots, x^n) = (x^{n-p+1}, x^{n-p+2}, \ldots, x^n).$$

In general we cannot assume that the last columns of the matrix give nonzero determinant. We know from Linear Algebra that there will be some p columns with this property. Let these columns be $j^1, j^2, \ldots, j^p$ with

$$M = (D_{j_k} g^i(a))_{i=1,\ldots,p\ k=1,\ldots,p}, \qquad det(M) \neq 0.$$

We rearrange the variables as follows: Let $m : \mathbb{R}^n \to \mathbb{R}^n$ be defined by (put the variables with superscript $j_k,\ k = 1, 2, \ldots, p$ in the last entries and order in whatever way you want the other variables)

$$m(x^1, x^2, \ldots, x^n) = (\ldots, x^{j_1}; x^{j_2}, \ldots, x^{j_p}).$$

Then $g \circ m$ is a function of the type discussed theorem **??**, so we can find a function $s : A \to \mathbb{R}^n$ which is bijective onto an open set V and s^{-1} is differentiable and

$$((g \circ m) \circ s)(x^1, x^2, \ldots, x^n) = (x^{n-p+1}; x^{n-p+2}, \ldots, x^n).$$

Then use $h = m \circ s$. $\square$

Example 1.13.

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x,y) = (xy, x^2+y^2) = (z,w)$$

$$\begin{pmatrix} \frac{dx}{dz} & \frac{dx}{dw} \\ \frac{dy}{dz} & \frac{dy}{dw} \end{pmatrix} = \begin{pmatrix} \frac{dz}{dx} & \frac{dz}{dy} \\ \frac{dw}{dx} & \frac{dw}{dy} \end{pmatrix}^{-1}$$

$$\begin{aligned} &z = xy, && y = \frac{z}{w} \\ &w = x^2+y^2 = x^2 + \frac{z^2}{x^2} \\ &wx^2 = x^4 + z^2 \\ &x^4 - wx^2 + z^2 && (*) \end{aligned}$$

$$x = g(z,w)$$

Use implicit differentiation on $(*)$ *with respect to* z*:*

$$4x^3\frac{dx}{dz} - w\cdot 2x\frac{dx}{dz} + 2z = 0$$

$$\frac{dx}{dz}(4x^3 - 2xw) = -2z$$

$$\frac{dx}{dz} = \frac{-2z}{4x^3-2xw} = \frac{-z}{x(2x^2-w)} = \frac{-y}{2x^2-w}$$

Valid for

$$\begin{gathered} 2x^2 - w \neq 0 \\ 2x^2 - (x^2+y^2) \neq 0 \\ x^2 - y^2 \neq 0 \\ \Leftrightarrow f'(x,y) \neq 0 \end{gathered}$$

$$\left.\begin{array}{ll} f(x,y) = 0 & f(a,b) = 0 \\ f(x,g(x)) = 0 & \textit{solve implicitly for } y \\ x \in \mathbb{R}^n,\ y \in \mathbb{R}^m & g : \mathbb{R}^n \to \mathbb{R}^m \\ f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m & \end{array}\right\} \quad \textit{Set up of implicit function theorem}$$

$$i = 1,\dots,m \quad f^i(x^1,\dots,x^n, g^1(x^1,\dots,x^n),\dots,g^m(x^1,\dots,x^n)) = 0$$

how to compute $D_j g^i$?

$$D_j g^i(\dots) = 0$$

$$D_1 f^i \cancel{\frac{dx^1}{dx^j}} + \cancel{\dots} + D_j f^i \underbrace{\frac{dx^j}{dx^j}}_{=1} + \cancel{\dots} + D_n f^i \cancel{\frac{dx^n}{dx^j}} + D_{n+1}f^i\frac{dg^1}{dx^j} + \dots + D_{n+m}f^i\frac{dg^m}{dx^j} = 0$$

$$\underbrace{D_{n+1}f^i\frac{dg^1}{dx^j} + \dots + D_{n+m}f^i\frac{dg^m}{dx^j}}_{m\ \textit{unknowns}} = -D_j f^i\frac{dx^j}{dx^j}$$

Check det of coefficents is $\neq 0$

$$\begin{bmatrix} D_{n+1}f^1 & \dots & D_{n+m}f^1 \\ \vdots & & \vdots \\ D_{n+1}f^m & \dots & D_{n+m}f^m \end{bmatrix} = M$$

2 Integration

2.1 Multiple integrals

$f : A \to \mathbb{R}$, A is a rectangle in $\mathbb{R}^n$ $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$
Recall a partition $\mathcal{P}$ of $[a, b]$ is a collection of of points: $t_0, \ldots, t_k$ with $a = t_0 < t_1 < \cdots < t_k = b$
A Partition of a rectangle $[a_1, b_1] \times \cdots \times [a_k, b_k]$ is a collection $\mathcal{P} = (P_1, \ldots, P_n)$ where P_i is a partition of $[a_i, b_i]$, $i = 1, \ldots, n$ Subrectangles $[s_{j-1}, s_j] \times [t_{m-1}, t_m]$ Let f be bounded on the rectangle $[a_1, b_1] \times \cdots \times [a_k, b_k]$

Definition 2.1. *Let f be bounded on the rectangle $[a_1, b_1] \times \cdots \times [a_k, b_k]$ and let S be subrectangle of the partition $\mathcal{P}$*

$$m_S(f) = \inf_{x \in S} f(x), \qquad M_S(f) = \sup_{x \in S} f(x)$$

Lower Riemann sum:

$$\mathcal{L}(f, \mathcal{P}) = \sum_S m_S(f).v(S)$$

where $v(s)$ is the volume of the subrectangle

$$S = [s_{l-1}, s_l] \times [t_{j-1}, t_j] \times \cdots \times [r_{k-1}, r_k]$$

$$v(S) = (s_{l-1} - s_l) \cdot (t_{j-1}, t_j) \cdots (r_{k-1}, r_k)$$

Upper Riemann sum:

$$\mathcal{U}(f, \mathcal{P}) = \sum_S M_S(f).v(S)$$

$$\mathcal{L}(f, \mathcal{P}) \leq \mathcal{U}(f, \mathcal{P})$$

Refinement: A refinement $\mathcal{P}'$ of the partition $\mathcal{P}$ is as follows. Given S a subrectangle of $\mathcal{P}'$, I can find a subrectangle T of $\mathcal{P}$ such that $S \subset T$ and $T = \cup_{S \subset T} S$, S for $\mathcal{P}'$

Lemma 2.1. *if $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then:*

$$\mathcal{L}(f, \mathcal{P}) \leq \mathcal{L}(f, \mathcal{P}') \tag{1}$$
$$\mathcal{U}(f, \mathcal{P}) \geq \mathcal{U}(f, \mathcal{P}') \tag{2}$$

Proof. of (1)
Let S be a subrectangle of $\mathcal{P}'$ and T a subrectangle of $\mathcal{P}$ such that $S \subset T$

$$\begin{aligned} m_S(f) &\geq M_T(f) \\ m_S(f)v(S) &\geq M_T(f)v(S) \qquad \text{(sum over all } S \subset T,\ S\ for\ \mathcal{P}') \\ \sum_{S \subset T} m_S(f)v(S) &\geq \sum_{S \subset T} m_T(f)v(S) = m_T(f)v(T) \\ \mathcal{L}(f, \mathcal{P}') = \sum_T \sum_{S \subset T} m_S(f)v(S) &\geq \sum_T m_T(f)v(T) = \mathcal{L}(f, \mathcal{P}) \end{aligned}$$

□

Lemma 2.2. *For any two partitions $\mathcal{P}$ and $\mathcal{P}'$ we have:*

$$\mathcal{L}(f,\mathcal{P}) \leq \mathcal{U}(f,\mathcal{P}')$$

Proof. Take $\mathcal{P}''$ a refinement of $\mathcal{P}$ and $\mathcal{P}'$:

$$\mathcal{L}(f,\mathcal{P}) \leq \mathcal{L}(f,\mathcal{P}'') \leq \mathcal{U}(f,\mathcal{P}'') \leq \mathcal{U}(f,\mathcal{P}')$$

□

Definition 2.2. *The lower Riemann integral*

$$\int_{A-} f = \sup_{\mathcal{P}} \mathcal{L}(f,\mathcal{P}), \quad (\mathcal{P} \textit{ partition of rectangle } A)$$

The upper Riemann integral

$$\bar{\int_{A}} f = \inf_{\mathcal{P}} \mathcal{U}(f,\mathcal{P})$$

f is called integrable if

$$\int_{A-} f = \bar{\int_{A}} f \quad \textit{and} \quad \int_{A} f = \int_{A-} f = \bar{\int_{A}} f$$

Theorem 2.3 (Riemann's Integrability Criterion)**.** *f is integrable over the rectangle $A \Leftrightarrow$ $\forall \epsilon > 0$, $\exists$ a partition $\mathcal{P}$ of A such that*

$$\mathcal{U}(f,\mathcal{P}) - \mathcal{L}(f,\mathcal{P}) < \epsilon$$

Proof. $(\Rightarrow)$

$$\begin{aligned}
&\inf_{\mathcal{P}}(\mathcal{U}(f,\mathcal{P}) - \mathcal{L}(f,\mathcal{P}) = 0 \\
\Leftrightarrow &\inf_{\mathcal{P}} \mathcal{U}(f,\mathcal{P}) - \sup_{\mathcal{P}} \mathcal{L}(f,\mathcal{P}) = 0 \\
\Leftrightarrow &\int_{A-} f = \bar{\int_{A}} f
\end{aligned}$$

$(\Leftarrow)$

Assume $\int_{A-} f = \bar{\int_{A}} f$, fix $\epsilon > 0$

$$\text{Since } \int_{A-} f = \sup_{\mathcal{P}} \mathcal{L}(f,\mathcal{P}), \quad \text{so} \quad \exists \mathcal{P}' \; s.t \int_{A-} f - \frac{\epsilon}{2} < \mathcal{L}(f,\mathcal{P}')$$

$$\text{Since } \overline{\int_A} f = \inf_{\mathcal{P}} \mathcal{U}(f, \mathcal{P}), \quad \text{so} \quad \exists \mathcal{P}' \; s.t \; \overline{\int_A} f + \frac{\epsilon}{2} > \mathcal{U}(f, \mathcal{P}')$$

Take $\mathcal{P}''$ a common refinement of $\mathcal{P}$ and $\mathcal{P}'$

$$\overline{\int_A} f + \frac{\epsilon}{2} \; > \; \mathcal{U}(f, \mathcal{P}'') \geq \mathcal{L}(f, \mathcal{P}'') \; > \underline{\int_A} f - \frac{\epsilon}{2}$$

So

$$\mathcal{U}(f, \mathcal{P}'') - \mathcal{L}(f, \mathcal{P}'') < \left(\cancel{\overline{\int_A} f} + \frac{\epsilon}{2} \right) - \left(\cancel{\underline{\int_A} f} - \frac{\epsilon}{2} \right) = \epsilon$$

$\square$

Example 2.1. *Non-Riemann integrable function* $f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x, y) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$$

$$\begin{aligned} m_S(f) &= 0 & M_S(f) &= 1 \\ \mathcal{L}(f, \mathcal{P}) &= 0 & \mathcal{U}(f, \mathcal{P}) &= 1 \end{aligned}$$

If $C \subset \mathbb{R}^n$, define the characteristic function of C to be

$$X_C(x) = \begin{cases} 1 & \text{if } x \in C \\ 0 & \text{if } x \notin C \end{cases}$$

If f is bounded on $\bar{C}$ and C is contained in a rectangle A, we define

$$\int_C f = \int_A f X_C$$

$f : [a, b] \times [c, d] \to \mathbb{R}$
Fix x and consider $g_x : [c, d] \to \mathbb{R}$

$$g_x(y) = f(x, y)$$

$$I(x) = \int_c^d g_x dy = \int_c^d f(x, y) dy$$

$$\int_a^b I(x) dx = \int_a^b \left(\int_c^d f(x, y) dy \right) dx \tag{1}$$

Fix y and define $h_y : [a, b] \to \mathbb{R}$

$$h_y(x) = f(x, y)$$

$$J(y) = \int_a^b h_y dx = \int_a^b f(x, y) dx$$

$$\int_c^d J(y) dy = \int_c^d \left(\int_a^b f(x, y) dx \right) dy \tag{2}$$

$(1) = (2)$

2.2 Fubini's theorem

Theorem 2.4. *Let A be a rectangle in $\mathbb{R}^n$ and let B be a rectangle in $\mathbb{R}^m$. $f : A \times B \to \mathbb{R}$ is intergrable. define:*

$$g_x : B \to \mathbb{R} \quad by \quad g_x = f(x,y), \qquad \forall y \in B,\ \forall x \in A$$

and let:

$$\left.\begin{aligned} \mathfrak{L}(x) &= \int_{B-} g_x = \int_{B-} f(x,y)dy \\ \mathfrak{U}(x) &= \int_{B}^{-} g_x = \int_{B}^{-} f(x,y)dy \end{aligned}\right\} \qquad exists\ \forall x \in A$$

Then $\mathfrak{L}(x)$ and $\mathfrak{U}$ are intergrable over A, and:

$$\int_A \mathfrak{L}(x)dx = \int_A \left(\int_{B-} f(x,y)dy \right) dx = \int_A \left(\int_B^{-} f(x,y)dy \right) dx = \int_A \mathfrak{U}(x)dx = \int_{A\times B} f$$

Proof. Let $\mathcal{P}_A$ be a partition of A, $\mathcal{P}_B$ be a partition of B. Let S_A a subrectangle of A, S_B a subrectangle of B. Then the rectangles $S_A \times S_B$ give a partition $\mathcal{P}$ of $A \times B$.
We will prove:

$$\mathcal{L}(f,\mathcal{P}) \underset{(1)}{\leq} \mathcal{L}(\mathfrak{L},\mathcal{P}_A) \underset{(2)}{\leq} \mathcal{U}(\mathfrak{L},\mathcal{P}_A) \underset{(3)}{\leq} \mathcal{U}(\mathfrak{U},\mathcal{P}_A) \underset{(4)}{\leq} \mathcal{U}(f,\mathcal{P})$$

Since f is integrable over $A \times B$, given $\epsilon > 0$ Riemann's integrability criterion given a partition $\mathcal{P}$ of $A \times B$, such that: $\mathcal{U}(f,\mathcal{P}) - \mathcal{L}(f,\mathcal{P}) < \epsilon$. Then $\mathcal{P}$ defines $\mathcal{P}_A$, $\mathcal{P}_B$ partitions of A, B respectively. By the inequality above: $\mathcal{U}(\mathfrak{L},\mathcal{P}_A) - \mathcal{L}(\mathfrak{L},\mathcal{P}_A) < \epsilon$. By reimann's integrability criterion, $\mathcal{L}$ is integrable over A, since:

$$\sup_{\mathcal{P}} \mathcal{L}(f,\mathcal{P}) = \inf_{\mathcal{P}} \mathcal{U}(f,\mathcal{P}) = \int_{A\times B} f \Rightarrow \int_A \mathfrak{L}(x)dx = \sup_{\mathcal{P}_A} \mathcal{L}(\mathfrak{L},\mathcal{P}_A) = \inf_{\mathcal{P}_A} \mathcal{U}(\mathfrak{L},\mathcal{P}_A) = \int_{A\times B} f$$

Works simularly with $\mathfrak{U}(x)$.
Side remark:

$$\begin{aligned} \mathcal{L}(f,\mathcal{P}) &\leq \mathcal{L}(\mathfrak{L},\mathcal{P}_A) \\ \sup_{\mathcal{P}} \mathcal{L}(f,\mathcal{P}) &\leq \sup_{\mathcal{P}_A} \mathcal{L}(\mathfrak{L},\mathcal{P}_A) \\ \mathcal{U}(\mathfrak{L},\mathcal{P}_A) &\leq \mathcal{U}(f,\mathcal{P}) \\ \inf_{\mathcal{P}_A} \mathcal{U}(\mathfrak{L},\mathcal{P}_A) &\leq \inf_{\mathcal{P}} \mathcal{L}(f,\mathcal{P}) \end{aligned}$$

(2) $\mathcal{L}(\mathfrak{L}, \mathcal{P}_A) \leq \mathcal{U}(\mathfrak{L}, \mathcal{P}_A)$

always true for a function $\mathfrak{L}$, partition $\mathcal{P}_A$

(3) $\mathcal{U}(\mathfrak{L}, \mathcal{P}_A) \leq \mathcal{U}(\mathfrak{U}, \mathcal{P}_A)$

$$\mathfrak{L}(x) = \int_{B\,-} f(x,y)dy,\ \mathfrak{U}(x) = \overline{\int_{B}} f(x,y)dy \Rightarrow \mathfrak{L}(x) \leq \mathfrak{U}(x)$$

$$\Rightarrow \mathcal{U}(\mathfrak{L}(x), \mathcal{P}_A) \leq \mathcal{U}(\mathfrak{U}(x), \mathcal{P}_A)$$

(4) is proved simularly to (1) so we only prove (1).

$$\mathcal{L}(f, \mathcal{P}) = \sum_S m_s(f)v(S) = \sum_{S_A,\, S_B} m_{S_A \times S_B}(f)v(S_A \times S_B) = \sum_{S_A} \left(\sum_{S_B} m_{S_A \times S_B}(f)v(S_B) \right) v(S_A)$$

Now, if $x \in S_A$, then clearly $m_{S_A \times S_B}(f) \leq m_{S_B}(g_x)$. Consequently, for $x \in S_A$ we have

$$\sum_{S_B} m_{S_A \times S_B}(f) \cdot v(S_B) \leq \sum_{S_B} m_{S_B}(g_x) \cdot v(S_B) \leq \int_{B\,-} g_x = \mathfrak{L}(x).$$

Therefore

$$\sum_{S_A} \left(\sum_{S_B} m_{S_A \times S_B}(f)v(S_B) \right) v(S_A) \leq \mathcal{L}(\mathfrak{L}, \mathcal{P}_A).$$

When the function is reimenn integral $\square$

2.3 Change of variables

Theorem 2.5. *Let $A \in \mathbb{R}^n$ be open, $g : A \to \mathbb{R}^n$ be injective and continuously differentiable with $det\ g'(x) \neq 0,\ \forall x \in A$. Let $f : g(A) \to \mathbb{R}$ be integrable. Then we have change of variables formula:*

$$\int_{g(A)} f = \int_A (f \circ g) \cdot |det\ g'(x)| dx$$

3 Differential Forms

3.1 Manifolds

Definition 3.1 (C^∞)**.** *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called a C^∞ function if all partial derivitives of all orders of all components exists and are continuous*

Definition 3.2 (Diffeomorphism)**.** *Let $U,\ V$ be open sets in $\mathbb{R}^n$. A C^∞ function $h : U \to V$ bijective and $h^{-1} : V \to U$, also a C^∞ function, is called a diffeomorphism from U to V*

Definition 3.3. *A set M is a K-dim manifold in $\mathbb{R}^n$ if the following condition (M) holds. For every $x \in M$:*
(M): There exisits two open sets U, V of $\mathbb{R}^n$, z $x \in U$ and a diffeomorphsim $h : U \to V$ such that:

$$h(U \cap M) = \{y \in V \ s.t. \ y^{k+1} = y^{k+2} = \cdots = y^n = 0\}$$

Theorem 3.1. *Let $A \to \mathbb{R}^n$ be open and let $g : A \to \mathbb{R}^p$ be a differentiable function such that $g'(x)$ has rank p on the set $g^{-1}(0)$. Then $g^{-1}(0)$ is an $n-p$ dimensional manifold in $\mathbb{R}^n$.*

Proof. It follows directly from theorem **??**. Let $x \in g^{-1}(0) = M$. We take $V = A$ in theorem**??** so that we can find a diffeomorphism $H : V \to U$, where U is open in $\mathbb{R}^n$ and

$$g \circ H(x^1, x^2, \ldots, x^n) = (x^{n-p+1}, x^{n-p+2}, \ldots, x^n)$$

Let $h = H^{-1} : U \to V$. We need to show that

$$h(U \cap M) = \{y \in V, \ y^{n-p+1} = y^{n-p+2} = \cdots = y^n = 0\}.$$

Let $y \in U \cap M$. Then $y \in g^{-1}(0)$, i.e. $g(y) = 0$. Since

$$h(g^1(0)) = H^{-1}(g^{-1}(0)) = (g \circ H)^{-1}(0) = \{y \in V, \ y^{n-p+1} = y^{n-p+2} = \cdots = y^n = 0\},$$

clearly we have for $y \in g^{-1}(0)$ that $h(y)$ has its last p coordinates zero. The converse is also obvious: if $z \in \{y \in V, \ y^{n-p+1} = y^{n-p+2} = \cdots = y^n = 0\}$ then set $y = H(z) \in U$ and $g(y) = g(H(z)) = (z^{n-p+1}, \ldots, z^n) = (0, \ldots, 0) \Rightarrow y \in g^{-1}(0) = M$ and $z = h(y) \in h(U \cap M)$. □

The following theorem gives the coordinate definition of a manifold M.

Theorem 3.2. *A subset M of $\mathbb{R}^n$ is a k-dimensional manifold iff for every point $x \in M$ the following holds:*
(C) There exists an open set $U \in \mathbb{R}^n$, $x \in U$ and an open set $W \subset \mathbb{R}^k$ and an injective differentiable map $f : W \to \mathbb{R}^n$ such that

(i) $f(W) = U \cap M$

(ii) *rank* $f'(y) = k \quad \forall y \in W$

(iii) $f^{-1} : f(W) \to W$ *is continuous.*

Proof. Lets assume that M is a manifold according to the deffnition (M). We choose the function $h : U \to V$ as in the definition. We define the set W and the function f as follows:

$$W = \{a \in \mathbb{R}^k, \ (a, 0) \in h(U \cap M)\}, \quad f : W \to \mathbb{R}^n, \quad f(a) = h^{-1}(a, 0).$$

Here $(a, 0)$ is the vector with the last $n-k$ coordinates equal to 0. Obviously $f(W) = U \cap M$, since

$$a \in W \Leftrightarrow (a, 0) \in h(U \cap M) \Leftrightarrow h^{-1}(a, 0) \in U \cap M \Leftrightarrow f(a) \in U \cap M.$$

We prove that W is open. For $a \in W$, we have:

$$(a, 0) \in h(U \cap M) \Leftrightarrow h^{-1}(a, 0) \in U \cap M \Rightarrow h^{-1}(a, 0) \in U.$$

Since h^{-1} is continuous, if b is sufficiently close to a, so that $(a, 0)$ and $(b, 0)$ are sufficiently close, we can deduce that $h^{-1}(b, 0)$ is close enough to $h^{-1}(a, 0)$. Because U open, if $h^{-1}(a, 0) \in$

U, then also $h^{-1}(b,0) \in U$. This gives $(b,0) \in h(U)$. Because $h(U \cap M)$ consists exactly of the points with last $n-k$ components equal to 0, $(b,0) \in h(U \cap M) \Leftrightarrow b \in W$. We immediately see from the definition of f and W that f^{-1} is continuous (it maps $h^{-1}(a,0)$ to a while h is continuous).
We prove that the rank of $f'(y)$ is k on W. For this we introduce another function

$$H : U \to \mathbb{R}^k, \quad H(z) = (h^1(z), \dots, h^k(z)),$$

i.e. H has the same first k coordinates as h (and ignores the last $n-k$). We have

$$H(f(y)) = H(h^{-1}(y,0)) = y, \quad y \in W.$$

Therefore, $H'(f(y)) \cdot f'(y) = I_{k \times k}$ or, in terms of linear transformations:

$$DH(f(y)) \circ Df(y) = Id_{\mathbb{R}^k}.$$

Because the composition is injective, $Df(y)$ is injective and the nullity plus rank theorem for $Df(y) : \mathbb{R}^k \to \mathbb{R}^n$ gives that the rank of $Df(y)$ is k.

The converse: Suppose that $f : W \to \mathbb{R}^n$ satisfies condition (C). We have $f'(y) \in M_{n \times k}$. By rearranging the coordinates in $\mathbb{R}^n$, we can assume that the rank of the first k rows of $f'(a)$ is k. This means

$$det(D_j f^i(a))_{i,j=1,\dots,k} \neq 0.$$

We define

$$g : W \times \mathbb{R}^{n-k} \to \mathbb{R}^n, \quad g(a,b) = f(a) + (0,b),$$

where $(0,b)$ has the first k coordinates 0. We have

$$g^i(a,b) = f^i(a), \quad i \leq k, \quad g^i(a,b) = f^i(a) + b^i, \quad i > k.$$

We compute its Jacobian matrix. For $i \leq k$

$$D_j g^i(a,b) = D_j f^i(a) \Rightarrow D_j g^i(a,b) = D_j f^i(a), \quad j \leq k,$$
$$and \quad D_j g^i(a,b) = 0, \quad j > k.$$

For $i > k$, however, we have

$$D_j g^i(a,b) = D_j f^i(a) + D_j b^i \Rightarrow D_j g^i(a,b) = \delta_{ij}, \quad j > k$$

while

$$D_j g^i(a,b) = D_j f^i(a), \quad j \leq k.$$

The Jacobian matrix is therefore in block form

$$g'(a,b) = \left(\begin{array}{c|c} D_j f^i(a)_{i,j=1,\dots,k} & 0 \\ \hline D_j f^i(a)^{i=k+1,\dots,n}_{j=1,\dots,k} & I_{(n-k)\times(n-k)} \end{array} \right)$$

The calculation of the determinant in block form (which can be considered as successive expansion on the last column) gives that $det\ g'(a,b) \neq 0$. By the inverse function theorem, there exists an open set V_1 with $(a,0) \in V_1$ and an open set V_2 containing $g(a,0) = f(a)$, such that $g : V_1 \to V_2$ has a differentiable inverse $h : V_2 \to V_1$. Then, since $f(W) = U \cap M$, we have for $(x,0) \in V_1$, $g(x,0) \in M \Leftrightarrow f(x) \in M$: This gives

$$V_2 \cap M = \{g(x,0),\ (x,0) \in V_1\}.$$
$$h(V_2 \cap M) = g^{-1}(V_2 \cap M) = g^{-1}(\{g(x,0),\ (x,0) \in V_1\}) = V_1 \cap (\mathbb{R}^k \times \{0\}).$$

$\square$

We also need the definition of manifold with boundary. While a k-dimensional manifold in $\mathbb{R}^n$ looks like a k-dim slice of $\mathbb{R}^n$, according to condition (M), for a manifold with boundary in $\mathbb{R}^n$, the part close to the boundary looks likes a half-slice of dimension k. To make this precise we define the half-space

$$\mathbb{H}^k = \{x \in \mathbb{R}^k,\ x^k \geq 0\}.$$

Then

$$h(U \cap M) = \{y \in V :\ y^k \geq 0,\ y^{k+1} = y^{k+2} = \cdots = y^n = 0\}$$

is the substitute for condition (M). More precisely:

Definition 3.4. *A subset M of $\mathbb{R}^n$ is a k-dimensional manifold with boundary if for every point x of M either condition (M) holds or (exclusive) the following condition holds:*
(M') There is an open set $U \to \mathbb{R}^n$ containing x, an open set V contained in $\mathbb{R}^n$ and a diffeomorphism $h : U \to V$ such that

$$h(U \cap M) = V \cap (\mathbb{H}^k \times \{0\}) = \{y \in V :\ y^k \geq 0,\ y^{k+1} = y^{k+2} = \cdots = y^n = 0\}.$$

Moreover, $h^k(x) = 0$.The set of points where condition (M') holds is called the boundary of M and is denoted by ∂M.

3.2 Linear Functionals

Definition 3.5. *Let $g^i : \mathbb{R}^n \to \mathbb{R}$ be a linear map, such a map is called a linear functional. The set of all linear functionals from $\mathbb{R}^n \to \mathbb{R}$ is called the dual space of $\mathbb{R}^n$, denoted $(\mathbb{R}^n)*$ let $g^1, \ldots, g^m$ be linear functionals $g^i : \mathbb{R}^n \to \mathbb{R}$, then I can combine them to get a map $g : \mathbb{R}^n \to \mathbb{R}^m$ by $g(x) = (g^1(x), \ldots, g^m(x))$*
$g : \mathbb{R}^n \to \mathbb{R}^m$ is linear such for $x, y \in \mathbb{R}^n,\ \lambda \in \mathbb{R}$

$$g(\lambda x + y) = \lambda g(x) + g(y)$$

this can be seen by

$$\begin{aligned} g(\lambda x + y) &= (g^1(\lambda x + y), \ldots, g^m(\lambda x + y) \\ &= (\lambda g(x)^1 + g^1(y), \ldots, \lambda g(x)^m + g^m(y)) \\ &= \lambda(g^1(x), \ldots, g^m(x)) + (g^1, \ldots, g^m) \end{aligned}$$

$[g^i]$ is the matrix representation of g^i
$[g^i] = (g^i_1, \ldots, g^i_n)$

$$[g]_{mxn} = \begin{pmatrix} g^1_1 & \cdots & g^1_n \\ \vdots & & \vdots \\ g^m_1 & \cdots & g^m_n \end{pmatrix}$$

Theorem 3.3. *$f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a iff f^i are differentiable at a, $i = 1, \ldots, m$ and $Df(a) = (Df^1, \ldots, Df^m(a))$*

Proof. assume f is differentiable at a we take the linear function $\Pi^i(x^1, \ldots, x^m) = x^i$ and compose it with f we get

$$f^i = \Pi^i \circ f$$

this is differentiable by chain rule since f and Π^i are differentiable $\forall i = 1, \ldots, m$

$$\Rightarrow Df^i = D\Pi^i(a) \cdot Df(a)$$

$D\Pi^i = \Pi^i$

$$\Rightarrow Df^i = \Pi^i(a) \cdot Df(a)$$

Now assume the all f^i are differentiable at a $\forall i = 1, \ldots, m$

$$\begin{aligned}
&f(a+h) - f(a) - (Df^1(a)(h), \ldots, Df^m(a)(h)) \\
&= (f^1(a+h), \ldots, f^m(a+h)) - (f^1, \ldots, f^m) - (Df^1 * (a)(h), \ldots, Df^m(a)(h)) \\
&= (f^1(a+h) - f^1(a) - df^1(a), \ldots, f^m(a+h) - f^m(a) - df^m(a))
\end{aligned}$$

So

$$\begin{aligned}
&\frac{|f(a+h) - f(a) - (Df^1(a)(h), \ldots, Df^m(a)(h))|}{|h|} \\
&\leq \frac{|f^1(a+h) - f^1(a) - df^1(a)|}{|h|}, \ldots, \frac{|f^m(a+h) - f^m(a) - df^m(a)|}{|h|} \to 0
\end{aligned}$$

$\square$

Remark. *If $T, S : \mathbb{R}^n \to \mathbb{R}^m$ are linear then $(T+S) : \mathbb{R}^n \to \mathbb{R}^m, (T+S)(x) = T(x) + S(x)$ is linear.*
If $\lambda \in \mathbb{R}$ then $(\lambda T) : \mathbb{R}^n \to \mathbb{R}^m, (\lambda T)(x) = \lambda \cdot T(x)$ is also linear.