Quick review for some linear algebra:

vector space: ...

- Linearly independent: $c_1\vec{v_1}+c_2\vec{v_2}+\cdots c_n\vec{v_n}=0$ only have trivial solution.
- $B=\{\vec{v}_\alpha \in V \mid \alpha\in A\}$ is a basis if i) $\text{span}(B)=V$, ii) B is linearly independent.

Thm 1: Every vector space has a bases.

Thm 2: Let r be a positive integer, if X spanned by set of r vectors. $X\le r$. ...

$\boxed{\text{Def 9.4:}}$ If V, W are $\mathbb{R}$-vector spaces, map: $A: V\to W$ is linear transformation if

$$A(\vec{x}_1+\vec{x}_2)=A\vec{x}_1+A\vec{x}_2$$

If $\exists\ B: W\to V$ which is linear such that: $B\circ A=$ identity map.

A is isomorphism.

Notation: $L(V,W)=\{\text{Linear map } A: V\to W\}$

$\boxed{\text{Thm}}$ Assume V is finite dimensional and let $A\in L(V)$ then following are equivalent.

1) A is injective 2) A is surjective 3) A is bijective.

$\boxed{\text{Def:}}$ Norm of a linear map. let $A\in L(\mathbb{R}^n,\mathbb{R}^m)$. $\|A\|=\sup\limits_{|\vec{x}|\le 1}|A\vec{x}|$

$\boxed{\text{Thm:}}$ Let $A\in (\mathbb{R}^n,\mathbb{R}^m)$. Then (a) $\|A\|<\infty$ (b) A is uniformly continuous.

(c) $\|\ \|$ is a norm of $L(\mathbb{R}^n,\mathbb{R}^m)$. $\Rightarrow$ $\|A\|=0 \Leftrightarrow A=0$

$\|cA\|=|c|\,\|A\|$ and $\|A+B\|\le\|A\|+\|B\|$

(d) $A\in L(\mathbb{R}^n,\mathbb{R}^m)$, $B\in L(\mathbb{R}^n,\mathbb{R}^k)$ then $\|BA\|\le\|B\|\|A\|$

$\boxed{\text{Thm 9.8}}$: Let Ω be the set of all invertible linear operators on $\mathbb{R}^n$.

a) If $A\in\Omega$, $B\in L(\mathbb{R}^n)$ and $\|B-A\|\cdot\|A^{-1}\|<1$, then $B\in\Omega$.

b) Ω is an open subset of $L(\mathbb{R}^n)$ and the mapping $A\to A^{-1}$ is continous on Ω.

PE: (a) $\alpha=\frac{1}{\|A^{-1}\|}$, $\beta\in\|B-A\|$, by assumption. $\beta<\alpha$. let $\vec{x}\in\mathbb{R}^n$ be arbitary.

then $\alpha|\vec{x}|=\alpha|A^{-1}A\vec{x}|\le\alpha\|A^{-1}\|\,|A\vec{x}|=\alpha\cdot\frac{1}{\alpha}\cdot|A\vec{x}|=|A\vec{x}|\le$

$$|A\vec{x}-B\vec{x}|+|B\vec{x}|=|(A-B)\vec{x}|+|B\vec{x}|\le\|A-B\|\,|\vec{x}|+|B\vec{x}|=\beta|\vec{x}|+|B\vec{x}|$$

$(\alpha-\beta)|\vec{x}| \le |B\vec{x}|$, in particular, $B\vec{x} \neq 0$ if $\vec{x} \neq \vec{0}$

Recall : if $C: \mathbb{R}^n \to \mathbb{R}^m$ is a linear map. then $\text{kernel } C(C) = \{\vec{x} \in \mathbb{R}^n : C\vec{x} = \vec{0} \in \mathbb{R}^m\}$.

C is injective $\Leftrightarrow$ kernel $C = \{0\}$. B is injective. any injective linear

$B: \mathbb{R}^n \to \mathbb{R}^n$ is bijective. $\to$ hence invertible. i.e. $B \in \Omega$

Small perturbation. operators still invertible (In particular. Consider A a matrix) on A

(b) Replace $\vec{x}$ by $B^{-1}\vec{y}$ in. $(\alpha-\beta)|B^{-1}\vec{y}| \le |BB^{-1}\vec{y}| = |\vec{y}|$, taking $|\vec{y}| \le 1$.

$\|B^{-1}\| \le \frac{1}{\alpha-\beta}$, finally, $B^{-1} - A^{-1} = B^{-1}(A-B)A^{-1}$, $\|B^{-1}-A^{-1}\| \le \|B^{-1}\|\,\|A-B\|\,\|A^{-1}\|$ ...

$\boxed{\text{Def : Matrices :}}$ Suppose $\{\vec{x}_1, \dots \vec{x}_n\}$, $\{\vec{y}_1, \vec{y}_2, \dots \vec{y}_m\}$ are bases of vector space X, Y

Then $\forall A \in L(X, Y)$: $A\vec{x}_j = \sum_{i=1}^{m} a_{ij} y_i$

$$[A]_{m\times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & & a_{mn} \end{bmatrix}, \qquad \underline{\underline{A\vec{x} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j \right) y_i}}.$$

$|A\vec{x}| = \sum_i \left(\sum_j a_{ij} c_j\right)^2 \le \sum_i \left(\sum_j a_{ij}^2 \cdot \sum_j c_j^2\right) = \sum_{ij} a_{ij} |\vec{x}|^2$. by def. $\|A\| \le \left\{\sum_{ij} a_{ij}^2\right\}^{\frac{1}{2}}$

Differentiation : Derive derivative of a function whose domain is $\underline{\mathbb{R}^n}$.

Recall when $n=1$, $f: (a, b) \to \mathbb{R}^1$. $f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$, $\underline{f(x+h) - f(x) = f'(x)h + r(h)}$

$\lim_{h\to 0} \frac{r(h)}{h} = 0$ $\Rightarrow$ Sum of a linear functions that takes h to $f'(x)\cdot h$

Note : Every real number $\in \mathbb{R}^1$ rise to a linear operator on $\mathbb{R}^1$, $\alpha \in L(\mathbb{R}^1)$

Now Consider $\vec{f} : (a, b) \subset \mathbb{R}^1 \to \mathbb{R}^m$. $\to$ $\vec{f}(x+h) - \vec{f}(x) = h\vec{y} + \vec{r}(h)$, $\underline{\lim_{h\to 0} \frac{\vec{r}(h)}{h} = \vec{0}}$

$h \in \mathbb{R}^1$, $h\vec{y} \in \mathbb{R}^m$. $\vec{y} \in L(\mathbb{R}^1, \mathbb{R}^m)$. hence $\vec{f}'(x) \in L(\mathbb{R}^1, \mathbb{R}^m)$

We cannot apply previous derivative definition : $\frac{\mathbb{R}^m}{\mathbb{R}^n} \Rightarrow$ makes no sense.

Def 9.11. Suppose open set $E \subset \mathbb{R}^n$. $\vec{f}: E \to \mathbb{R}^m$. $\exists A \in L(\mathbb{R}^n, \mathbb{R}^m)$

$\lim_{\vec{h}\to 0} \frac{|\vec{f}(\vec{x}+\vec{h}) - \vec{f}(\vec{x}) - A\vec{h}|}{|\vec{h}|}$, then $\vec{f}$ is differentiable at $\vec{x}$, $\boxed{\vec{f}'(\vec{x}) = A}$

$\boxed{\text{Thm 9.12}}$ Suppose E and $\vec{f}$ are with above definition, $\vec{x} \in E$.

$A = A_1$. $A = A_2$. $A_1 = A_2$ the differentiation is unique.

Pf: Suppose $|(A_1 - A_2)\vec{h}| = |\vec{f}(\vec{x}_0+\vec{h}) - \vec{f}(\vec{x}_0) - A_2 h - (\vec{f}(\vec{x}_0+\vec{h}) - \vec{f}(\vec{x}_0) - A_1 h)|$

$\leq |\vec{f}(\vec{x}_0+\vec{h}) - \vec{f}(\vec{x}_0) - A_2\vec{h}| + |\vec{f}(\vec{x}_0+\vec{h}) - \vec{f}(\vec{x}_0) - A_1\vec{h}|$, divide by $|\vec{h}|$.

$\lim_{h \to \vec{0}} \frac{|(A_1 - A_2)\vec{h}|}{|\vec{h}|} = 0$, if $\vec{h}_0 \in \mathbb{R}^n$, $\vec{h}_0 \neq \vec{0}$. take $t \in \mathbb{R}^1$. $\lim_{t \to 0} \frac{|(A_1 - A_2)t\vec{h}_0|}{|t\vec{h}_0|} = 0$

Linearity of $A_1 - A_2$ (linear operator) shows that Independent of t. $\Rightarrow A_1 - A_2 = 0$.

$A_1 = A_2$. as desired.

$\boxed{\text{Remark 1:}}$ If $\vec{f}$ is differentiable in E, this gives $\boxed{\vec{f}': E \to L(\mathbb{R}^n, \mathbb{R}^m)}$

Remark 2: ..

Now introduce couple of propositions we know when $(m = n = 1)$.

$\boxed{\text{Proposition 1:}}$ $\vec{f}: E \to \mathbb{R}^m$ is differentiable at $\vec{x}_0 \in E \subset \mathbb{R}^n$, then it is continuous at $\vec{x}_0$

Pf: put $\vec{r}(\vec{h}) = \vec{f}(\vec{x}_0+\vec{h}) - \vec{f}(\vec{x}_0) - \vec{f}'(\vec{x}_0)\vec{h}$, then $\lim_{\vec{h} \to \vec{0}} \frac{|\vec{r}(\vec{h})|}{|\vec{h}|} = \vec{0}_{\mathbb{R}^m}$

In particular $\lim_{\vec{h} \to \vec{0}} \vec{r}(\vec{h}) = 0$ (proved above).

Since Linear maps are continuous. $\lim_{\vec{h} \to \vec{0}} \vec{f}'(\vec{x}_0)\cdot\vec{h} = \vec{0}_{\mathbb{R}^m}$. ($\vec{h} \in \mathbb{R}^n$. $\vec{f}' \in L(\mathbb{R}^n, \mathbb{R}^m)$)

$\underline{\lim_{\vec{h} \to \vec{0}} |\vec{f}(\vec{x}_0+\vec{h}) - \vec{f}(\vec{x}_0)|} = \lim_{\vec{h} \to \vec{0}} |\vec{r}(\vec{h}) + \vec{f}'(\vec{x}_0)\vec{h}| \leq \lim_{\vec{h} \to \vec{0}} |\vec{r}(\vec{h})| + \lim_{\vec{h} \to \vec{0}} |\vec{f}'(\vec{x}_0)\cdot\vec{h}| = \underline{\vec{0}}$

Thus we proved continuity.

$\boxed{\text{Proposition 2:}}$ Let $A: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $\forall \vec{x} \in \mathbb{R}^n$. $A'(\vec{x}) = A$

more general: $\vec{f}: \mathbb{R}^n \to \mathbb{R}^m$. $\vec{f}(\vec{x}) = \vec{v} + A\vec{x}$, then $\vec{f}'(\vec{x}) = A$.

Pf: The second assertion implies first. (take $\vec{v} = \vec{0}$).

$\vec{f}(\vec{x}+\vec{h}) - \vec{f}(\vec{x}) = A(\vec{x}+\vec{h}) - A\vec{x} = A\vec{h}$ (by linearity)

$\lim_{\vec{h} \to \vec{0}} \frac{|\vec{f}(\vec{x}+\vec{h}) - \vec{f}(\vec{x}) - A\vec{h}|}{|\vec{h}|} = \lim_{h \to \vec{0}} \frac{|A\vec{h} - A\vec{h}|}{|\vec{h}|} = \vec{0} \Rightarrow$ differential of $\vec{f}$ at

$\vec{x}$ is $\underline{A}$

Note that A is on the left-side. not right.

The several variable chain rule:

Recall for single variable: $\underline{\underline{(g\circ f)'(x) = g'(f(x_0))\cdot f'(x_0)}}$

Let $E\subset\mathbb{R}^n$ be open, $\vec{x}_0\in E$, $\vec{f}: E\to\mathbb{R}^m$, let $\Omega\subseteq\mathbb{R}^m$, $\vec{f}(E)\subseteq\Omega$. let $\vec{g}:\Omega\to\mathbb{R}^k$.

$\underline{\underline{\vec{g}\circ\vec{f}: E\to\mathbb{R}^k}}$ (Composition)

$\boxed{\text{Thm 9.15}}$ Suppose $\vec{f}$ is differentiable at $\vec{x}_0$ and $\vec{g}$ differentiable at $\vec{f}(\vec{x}_0)$, then $\vec{g}\circ\vec{f}$ is differentiable at $\vec{x}_0$ and $\underline{(\vec{g}\circ\vec{f})'(\vec{x}_0) = \vec{g}'(\vec{f}(\vec{x}_0))\,\vec{f}'(\vec{x}_0)}$

Comment: The right-hand side. $\vec{g}'(\vec{f}(\vec{x}_0))\in L(\mathbb{R}^m,\mathbb{R}^k)$ $\vec{f}'(\vec{x}_0)\in L(\mathbb{R}^n,\mathbb{R}^m)$ product should be composite linear map: $\underline{\underline{L(\mathbb{R}^n,\mathbb{R}^k)}}$

In diagrams:

$$\begin{tikzcd}
\vec{x}_0 \arrow[r, mapsto] & \vec{f}(x_0) \\
E\subset\mathbb{R}^n \arrow[r, "\vec{f}"] \arrow[dr, "\vec{g}\circ\vec{f}"'] & \Omega\subset\mathbb{R}^m \arrow[d, "\vec{g}"] \\
& \mathbb{R}^k
\end{tikzcd}$$

Pf: Put $\vec{y}_0=\vec{f}(\vec{x}_0)$. $A=\vec{f}'(x_0)$. $B=\vec{g}'(\vec{y}_0)=\vec{g}'(\vec{f}(\vec{x}_0))$, $\vec{F}=\vec{g}\circ\vec{f}: E\to\mathbb{R}^k$.

We need to show that: $\displaystyle\lim_{\vec{h}\to\vec{0}}\frac{|\vec{F}(\vec{x}_0+\vec{h})-\vec{F}(\vec{x}_0)-BA\vec{h}|}{|\vec{h}|}=0$. for all $\vec{h}\in\mathbb{R}^n$.

put $\underline{\vec{u}(\vec{h})=\vec{f}(\vec{x}_0+\vec{h})-\vec{f}(x_0)-A\vec{h}}$, Similarly, for all $\vec{k}\in\mathbb{R}^m$, $\vec{g}(\vec{y}_0+\vec{k})$ is defined.

put $\underline{\vec{v}(\vec{k})=\vec{g}(\vec{y}_0+\vec{k})-\vec{g}(\vec{y}_0)-B\vec{k}}$. In particular, these are defined for all $\vec{h},\vec{k}$ with $|\vec{h}|, |\vec{k}|$ sufficiently small! E/Ω are open sets.

Also put $\boxed{\varepsilon(\vec{h})=\frac{|\vec{u}(\vec{h})|}{|\vec{h}|}}$ $\boxed{\eta(\vec{k})=\frac{|\vec{v}(\vec{k})|}{|\vec{k}|}}$.

Now: differentiability of $\vec{f}$ at $\vec{x}_0$ is $\displaystyle\lim_{\vec{h}\to\vec{0}}\varepsilon(\vec{h})=0$. Similarly for $\vec{g}$: $(\vec{f}(\vec{x}_0))$

$\displaystyle\lim_{\vec{k}\to\vec{0}}\eta(\vec{k})=\vec{0}$, By proposition 1. $\vec{f}$ is continuous at $\vec{x}_0$. If $|\vec{h}|$ sufficiently small. $\vec{v}(\vec{k})$ defined at $\vec{k}=\vec{f}(\vec{x}_0+\vec{h})-\vec{f}(x_0)$, for $\vec{k}$, we have:

$|\vec{k}|=|\vec{u}(\vec{h})+A\vec{h}|$ (previous definition) $\le(\varepsilon(\vec{h})+\|A\|)|\vec{h}|$, $|\vec{h}|\to 0$ implies $|\vec{k}|\to 0$, Now. $\vec{F}(\vec{x}_0+\vec{h})-\vec{F}(\vec{x}_0)-BA\vec{h}=\vec{g}(\vec{y}_0+\vec{k})-\vec{g}(\vec{y}_0)-BA\vec{h}=$

$=\vec{v}(\vec{k})+B\vec{k}-BA\vec{h}=\vec{v}(\vec{k})+B(\vec{k}-A\vec{h})=\vec{v}(\vec{k})+B(\vec{u}(\vec{k})+A\vec{h}-A\vec{h})$

$=\vec{v}(\vec{k})+B\,\vec{u}(\vec{h})$

$$\frac{|\vec{F}(\vec{x}_0+\vec{h}) - \vec{F}(\vec{x}_0) - BA\vec{h}|}{|\vec{h}|} = \frac{|\vec{v}(\vec{k}) + B\vec{u}(\vec{h})|}{|\vec{h}|} \le \frac{\|B\|\,|\vec{u}(\vec{h})|}{|\vec{h}|} + \frac{|\vec{v}(\vec{k})|}{|h|}$$

$$\le \frac{\|B\|\,|\vec{u}(\vec{h})|}{|\vec{h}|} + \frac{\frac{|\vec{v}(\vec{k})|}{|\vec{k}|}}{\|A\|+\varepsilon(\vec{h})} = \|B\|\,\varepsilon(\vec{h}) + (\|A\|+\varepsilon(\vec{h}))\,\eta(\vec{k}) \quad \left(\lim_{h\to 0} \cdots \underline{0}\right)$$

※

Partial Derivatives :

$\vec{f}$ has components : $f_i : E \to \mathbb{R}$, $i=1,2,3\ldots m$, $\vec{f}(\vec{x}) = (f_1(\vec{x}), f_2(\vec{x}) \ldots f_m(\vec{x}))$

Alternatively, if $\{\vec{u}_1, \vec{u}_2, \ldots \vec{u}_m\}$ is a standard basis of $\mathbb{R}^m$, then the components f_i. $\vec{f}(\vec{x}) = \sum_{i=1}^{m} f_i(\vec{x})\,\vec{u}_i$, let $\{\vec{e}_1, \vec{e}_2 \ldots \vec{e}_n\}$ denotes the standard basis for $\mathbb{R}^n$.

Def : If $\vec{x}_0 \in E$, $i \in \{1,\ldots m\}$, $j \in \{1,\cdots n\}$. j^{th} partial derivative of f_i at $\vec{x}_0$ f_i w.r.t. x_j at $\vec{x}_0$ = $\boxed{(D_j f_i)(\vec{x}_0) = \frac{\partial f_i}{\partial x_j}(\vec{x}_0) = \lim_{t\to 0} \frac{f_i(\vec{x}_0 + t\vec{e}_j) - f_i(\vec{x}_0)}{t}}$

If the limit exists $1\le i\le m$, $1\le j\le n$. $\vec{x}\in E$. we can say f has partial derivative in E.

Thm 9.17 : If $\vec{f} : E \to \mathbb{R}^m$ is differentiable at $\vec{x}_0 \in E$, then the partial deri.. exist and $(D_j f_i)(\vec{x}_0)$ exist. and $\vec{f}'(\vec{x}_0)\,\vec{e}_j = \sum_{i=1}^{m} (D_j f_i)(\vec{x}_0)\,\vec{u}_i = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(\vec{x}_0)\,\vec{u}_i$, for each $1\le j\le n$

linear transform.

This equation is saying that $[\vec{f}'(x_0)] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{x}_0) \\ \vdots & & \\ \frac{\partial f_2}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_2}{\partial x_n}(\vec{x}_0) \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1}(\vec{x}_0) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{x}_0) \end{bmatrix}$

Pf : Since $\vec{f}$ differentiable at $\vec{x}_0$, $\vec{f}(\vec{x}_0+\vec{h}) - \vec{f}(\vec{x}_0) = \vec{f}'(\vec{x}_0)\vec{h} + \vec{r}(\vec{h})$, with $\lim_{\vec{h}\to\vec{0}} \frac{|\vec{r}(\vec{h})|}{|\vec{h}|} = 0$, fix $1\le j\le n$. and take $\vec{h} = t\vec{e}_j$, so that $\vec{f}'(\vec{x}_0)(t\vec{e}_j) = t\vec{f}'(\vec{x}_0)\vec{e}_j$

Then $\lim_{t\to 0} \frac{\vec{f}(\vec{x}_0 + t\vec{e}_j) - f(\vec{x}_0)}{t} = \frac{t\vec{f}'(x)\vec{e}_j + \vec{r}(t\vec{e}_j)}{t} = \lim_{t\to 0} \vec{f}'(x)\,\vec{e}_j + \lim_{t\to 0} \frac{\vec{r}(t\vec{e}_j)}{t}$

$= \vec{f}'(x_0)\,\vec{e}_j$, $\quad \vec{f}'(\vec{x}_0)\,\vec{e}_j = \lim_{t\to 0} \frac{\vec{f}(\vec{x}_0 + t\vec{e}_j) - \vec{f}(\vec{x}_0)}{t} = \lim_{t\to 0} \sum_{i=1}^{m} \frac{f_i \cdots}{\cdots}$

$= \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(\vec{x}_0)\,\vec{u}_i.$

Def: Let $\vec{v} \in \mathbb{R}^n$ be a unit vector, $|\vec{v}|=1$. $E \subset \mathbb{R}^n$ be open set. $\vec{x}_0 \in E$.
$\vec{f}: E \to \mathbb{R}^m$. the directional of $\vec{f}$ at $\vec{x}_0$ w.r.t $\vec{u}$ is:

$$(D_{\vec{v}}\vec{f})(\vec{x}_0) = \lim_{t\to 0} \frac{\vec{f}(\vec{x}_0 + t\vec{v}) - \vec{f}(\vec{x}_0)}{t} \in \mathbb{R}^m$$ [Directional derivative]

$$(D_{\vec{v}}\vec{f})(\vec{x}_0) = \sum_{i=1}^{m} (D_v f_i)(\vec{x}_0)\, \vec{u}_i$$

$$\nabla f(x_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}_0)\, \vec{e}_i = \left(\frac{\partial f}{\partial x_1}, \cdots \frac{\partial f}{\partial x_n}\right)$$ ⇒ gradient vector.

$\boxed{\text{Proposition: } (D_{\vec{v}} f)(\vec{x}_0) = (\nabla f)(\vec{x}_0)\cdot\vec{v}}$ dot product of 2 vectors in $\mathbb{R}^n$.

Pf: Consider $\gamma: \mathbb{R} \to \mathbb{R}^n$, $\gamma(t) = \vec{x}_0 + t\vec{v}$ → the image of $\gamma(t)$ is parallel to $\vec{v}$ and passes through point $\vec{x}_0$, $\gamma(0) = \vec{x}_0$. Let $g = f\circ\gamma$, $g: \mathbb{R} \to \mathbb{R}$.
$g(t) = f(\gamma(t))$. (g is f restricted to this line). By chain rule. $g'(0) = f'(\gamma(0))\cdot\gamma'(0)$

$g'(0) = f'(\vec{x}_0)\cdot\gamma'(0)$, Composition of $\gamma'(0) = L(\mathbb{R}, \mathbb{R}^n)$, $f'(\vec{x}_0) \in L(\mathbb{R}^n, \mathbb{R})$

Then the result should be: $L(\mathbb{R}, \mathbb{R}) \sim \mathbb{R}$

If we take standard bases $\{1\}$ of $\mathbb{R}$ and $\{\vec{e}_1, \dots \vec{e}_n\}$ of $\mathbb{R}^n$. the corresponding matrices are: $[\gamma'(0)] = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, $\vec{v} = \sum_{i=1}^{n} v_i \vec{e}_i$

$$[f'(\vec{x}_0)] = \left[\frac{\partial f}{\partial x_1}(\vec{x}_0), \cdots \frac{\partial f}{\partial x_n}(\vec{x}_0)\right] = \nabla f(\vec{x}_0)$$

$= g'(0)$.

The composition has matrix: $[f'(\vec{x}_0)][\gamma'(0)]$ = (composition of LT) = $\nabla f(\vec{x}_0)\cdot\vec{v}$.

Now on the other hand: $g'(0) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = \lim_{t\to\infty} \frac{f(\gamma(t)) - f(\gamma(0))}{t} =$

$\lim_{t\to 0} \frac{f(\vec{x}_0 + t\vec{v}) - f(\vec{x}_0)}{t} = (D_{\vec{v}} f)(\vec{x}_0) \in \mathbb{R}$. $(m=1)$. Thus

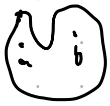
$\nabla f(\vec{x}_0)\cdot\vec{v} = (D_{\vec{v}} f)(\vec{x}_0)$ attains its maximum $\vec{v} = k\nabla f(\vec{x}_0)$ (scalar multiple).

$\boxed{\text{Rmark}}$: $\vec{v} = \sum v_i \vec{e}_i$. $(D_{\vec{v}} f)(\vec{x}_0) = \sum_{i=1}^{n} (D_i f)(\vec{x}) v_i$ → terms of partial derivative.

$\boxed{\text{Thm } 9.19}$: Let $E \subseteq \mathbb{R}^n$ be a convex open subset. Let $\vec{f}: E \to \mathbb{R}^n$ be differentiable in E. Suppose $\exists M \geq 0$ s.t. $\|\vec{f}'(\vec{x})\| \leq M$, $\forall \vec{x} \in E$. then $|\vec{f}(\vec{b}) - \vec{f}(\vec{a})| \leq M|\vec{b} - \vec{a}|$

This is weak version of the MVT for vector valued function

Pf: Set E is convex if $\forall \vec{a}, \vec{b} \in E$, the entire segment is contained in E.

 $\to$ not convex. Now we parametrized the segment: $\gamma : [0,1] \to \mathbb{R}^n$, $\gamma(t) = (1-t)\vec{a} + t\vec{b}$

Now consider $\vec{g} : [0,1] \to \mathbb{R}^n$, $\vec{g} = \vec{f} \circ \gamma$, $\vec{g}(t) = \vec{f}(\gamma(t))$, by chain rule:

$\vec{g}'(t) = \vec{f}'(\gamma(t)) \cdot \gamma'(t) = \vec{f}'(\gamma(t)) \cdot (\vec{b} - \vec{a})$

$|\vec{g}'(t)| \leq \|\vec{f}'(\gamma(t))\| \cdot |\vec{b} - \vec{a}| \leq M |\vec{b} - \vec{a}|$, by Thm 5.19. $\forall t$.

$|\vec{g}(1) - \vec{g}(0)| \leq M |\vec{b} - \vec{a}|$. $\vec{g}(1) = \vec{b}$, $\vec{g}(0) = a$ $\Rightarrow$ we finish the proof.

$\boxed{\text{Corollary}}$: If E is convex, and $\vec{f}'(\vec{x}_0) = 0$, $\forall \vec{x} \in E$, then $\vec{f}$ is constant.

$\forall \vec{a}, \vec{b}$, $|\vec{f}(\vec{b}) - \vec{f}(\vec{a})| \leq 0 = 0$. hence $\vec{f}$ is constant. (Apply $M = 0$ from above).

Now we've seen that $\vec{f}$ differentiable at $\vec{x}_0$ $\Rightarrow$ all partial derivative exists.

However the converse is in general not true. (Cannot guarantee continuity), need additional assumption in order to obtain a converse.

$\boxed{\text{Def } 9.20}$: Let $E \subset \mathbb{R}^n$, be open and let $\vec{f} : E \to \mathbb{R}^m$ be differentiable in E,

We say $\vec{f}$ is continously differentiable in E. if $\vec{f}' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$

is continous. Here think of $L(\mathbb{R}^n, \mathbb{R}^m)$ as the metric space with distance $d(A,B)$

$= \|A - B\|$. It means $\forall \vec{x} \in E$, $\forall \varepsilon > 0$. $\exists \delta > 0$. s.t. $\vec{y} \in E$, $|\vec{x} - \vec{y}| < \delta$,

$\Rightarrow \|\vec{f}'(\vec{y}) - \vec{f}'(\vec{x})\| < \varepsilon$. We write $C'(E, \mathbb{R}^m) = \{ \vec{f} : E \to \mathbb{R}^m \mid f \text{ is continously diff..} \}$

$\boxed{\text{Thm } 9.21}$: $\vec{f} \in C'(E, \mathbb{R}^m)$. $\Leftrightarrow$ all partial derivatives $D_j f_i = \dfrac{\partial f_i}{\partial x_j}$ exist and are continous in E.

Pf: Let $\{\vec{e}_1, \dots \vec{e}_n\}$ be the standard basis of $\mathbb{R}^n$. $\{u_1, u_2 \dots u_m\}$ be the standard basis in $\mathbb{R}^m$. $\vec{f}(\vec{x}) = \sum_{i=1}^{m} f_i(\vec{x}) \cdot \vec{u}_i$

"$\Rightarrow$" Assume that $\vec{f}$ is C' mapping, then by previous all $D_j f_i$ exists, we need to show they're continous. $\boxed{\text{by previous definition} : (\vec{f}'(\vec{x}) \vec{e}_j) \cdot \vec{u}_i = D_j f_i(\vec{x})}$

use of this is important!!

$$|D_j f_i(\vec{x}) - D_j f_i(\vec{y})| = |((\vec{f}'(\vec{x}) - \vec{f}'(\vec{y}))\vec{e}_j)\cdot\vec{u}_i| \le |(\vec{f}'(\vec{x}) - \vec{f}'(\vec{y}))\vec{e}_j|\,\underbrace{|\vec{u}_i|}_{1}$$

$$\le |\vec{f}'(\vec{x}) - \vec{f}'(\vec{y})|\,\underbrace{|\vec{e}_j|}_{1} = \|\vec{f}'(\vec{x}) - \vec{f}'(\vec{y})\|, \quad \text{fix } \vec{x} \text{ and } \varepsilon > 0.$$

Choose δ. s.t. $|\vec{x}-\vec{y}| < \delta \Rightarrow \|\vec{f}'(\vec{x}) - \vec{f}'(\vec{y})\| < \varepsilon$. (by our assumption).

Hence. $|D_j f_i(\vec{x}) - D_j f_i(\vec{y})| < \varepsilon$. , we proved the continuity.

$\boxed{\text{"}\Leftarrow\text{" For converse}}$ divide this into 2 steps $\quad f_i : E \to \mathbb{R}$

$\boxed{\text{Step 1: (Reduction to } m=1)}$. Suppose each component f_i is differentiable in E and each $f_i' : E \to L(\mathbb{R}^n, \mathbb{R})$ is continuous in E.. Fix $\vec{x} \in E$ and define a linear transformation $A_{\vec{x}} \in L(\mathbb{R}^n, \mathbb{R}^m)$. by $(*)$ $A_{\vec{x}}\cdot\vec{v} = \sum_{i=1}^{m}(f_i'(\vec{x})\vec{v})\cdot\vec{u}_i$

$(*)$. $\boxed{A_{\vec{x}}\vec{v} = \sum_{i=1}^{m}(f_i'(\vec{x})\vec{v})\vec{u}_i}$, $\vec{v} \in \mathbb{R}^n$. , then for $|\vec{h}|$ sufficiently small.

$$\frac{|\vec{f}(\vec{x}+\vec{h}) - \vec{f}(\vec{x}) - A_{\vec{x}}\vec{h}|}{|\vec{h}|} = \frac{|\sum_{i=1}^{m}(f_i(\vec{x}+\vec{h}) - f_i(\vec{x}) - f_i'(\vec{x})\vec{h})\cdot\vec{u}_i|}{|\vec{h}|} \le$$

Schwarz - inequality.

$$\sum_{i=1}^{m}\frac{|f_i(\vec{x}+\vec{h}) - f_i(\vec{x}) - f_i'(\vec{x})\vec{h}|}{|\vec{h}|}\cdot|\vec{u}_i|,$$ by assumption each term in the sum $\to 0$

as $\vec{h} \to \vec{0}$., hence $\vec{f}$ is differentiable at $\vec{x}$ and $\underline{\vec{f}'(\vec{x}) = A_{\vec{x}}}$ (Def of derivative).

$$|(A_{\vec{x}} - A_{\vec{y}})\vec{v}| = |\sum_{i=1}^{m}((f_i'(\vec{x}) - f_i'(\vec{y}))\vec{v})\vec{u}_i| \le |\sum_{i=1}^{m}(f_i(\vec{x}) - f_i(\vec{y}))\vec{v}|\cdot|\vec{u}_i|$$

$$= |\sum_{i=1}^{m}(f_i(\vec{x}) - f_i(\vec{y}))\cdot\vec{v}| \le \sum_{i=1}^{m}|f_i(\vec{x}) - f_i(\vec{y})\cdot\vec{v}|$$

(Schwarz))

$$\|A_{\vec{x}} - A_{\vec{y}}\| \le \sum_{i=1}^{m}|f_i(\vec{x}) - f_i(\vec{y})|.$$ this gives the continuity of $\vec{f}'$

$\boxed{\text{Step 2:}}$ Now take $f : E \to \mathbb{R}$. , and show f is differentiable (hence continous) in E if all $\frac{\partial f}{\partial x_j}$ exists are continous in E.

Now our next goal is the Inverse function theorem. Which is the most useful one.

First let's state the theorem. Inverse Function theorem :

$\boxed{\text{Thm 9.24:}}$ Suppose $\vec{f}$ is $\mathcal{C}'(E, \mathbb{R}^n)$, $\vec{f}'(\vec{a})$ is invertible for some $\vec{a} \in E$. $\vec{b} = \vec{f}(\vec{a})$

Then a) There exists open sets U and V in $\mathbb{R}^n$ s.t. $\vec{a} \in U$, $\vec{b} \in V$. $\vec{f}$ is one-to-one on U. and $\vec{f}(U) = V$

b) If $\vec{g}$ is the inverse of $\vec{f}$, defined in V, by $\vec{g}(\vec{f}(x)) = \vec{x}$ ($\vec{x} \in U$).

then $\vec{g} \in \mathcal{C}'(V)$

$\boxed{\text{Corollary:}}$ Let $E \subseteq \mathbb{R}^n$. be open. and $\vec{f}: E \to \mathbb{R}^n$, $\vec{f} \in \mathcal{C}'(E, \mathbb{R}^n)$, and if $\vec{f}'(\vec{x})$ is invertible for $\forall \vec{x} \in E$. then $\vec{f}(W)$ is an open subset of $\mathbb{R}^n$ for all open set $W \subset E$.

$\boxed{\text{Remark}}$: ① A function $\vec{f}$ satisfies the inverse function theorem is $\Rightarrow$ locally invertible near point $\vec{a}$: derivative (Jacobian) invertible $\Rightarrow$ function itself is invertible.

② If $m=n=1$. then (i) is easy to prove.

We have $a \in E \subseteq \mathbb{R}$, E open and $f: E \to \mathbb{R}$ differentiable on E, with continuous $f': E \to \mathbb{R}$; $b = f(a)$. Invertibility of the derivative at a simply means $f'(a) \neq 0$. Because f' is continuous, there is an open interval $I \subseteq E$, $a \in I$ such that either $f'(x) > 0$ $\forall x \in I$, or $f'(x) < 0$, $\forall x \in I$ (depending on the sign of $f'(a)$). Hence f is strictly monotonic on I, and in particular it is injective on I. The image $f(I)$ is an interval J containing b, which we may assume to be open (if not, replace J by its interior, and I by $f^{-1}(\mathring{J})$)

③ Let's understand what this theorem says when writing down its coordinate:

Let $\vec{x} = (x_1, \dots x_n)$. $\vec{a} = (a_1, \dots a_n)$

$(y_1, \dots y_n) = \vec{f}(\vec{x}) = (f_1(x_1 \dots x_n), \dots f_n(x_1 \dots x_n))$

$\vec{b} = (b_1, \dots b_n)$ with $b_i = f_i(a_1, \dots a_n)$

1) All partial derivatives exists and are continous. and the Jacobian:

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a_1 \dots a_n) & \cdots & \frac{\partial f_1}{\partial x_n}(a_1 - a_n) \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1}(a_1, \dots a_n) & \cdots & \frac{\partial f_n}{\partial x_n}(a_1, \dots a_n) \end{bmatrix} \in \mathrm{Mat}_{n \times n}(\mathbb{R}) \quad \text{is invertible.}$$

$\Rightarrow$ its determinant is not 0.

then the system of equations:

$$\begin{cases} y_1 = f_1(x_1 \dots x_n) \\ \vdots \\ y_n = f_n(x_1 \dots x_n) \end{cases}$$

for $(x_1 \dots x_n)$, $(y_1 \dots y_n)$ in a small neighborhood. has unique solution.

Moreover. the functions $g_1 \dots g_n$ all have continous partial derivatives in V.

④ $\vec{f}$ need not be globally invertible Ex: $\vec{f}(x,y) = (e^x \cos y, e^x \sin y)$.

see exercise 17.

⑤ if $\vec{f}$ is invertible. (locally or globally). the derivative at $\vec{a}$ need not be invertible.

To prove the Inverse function theorem. lets first introduce contraction principle:

$\boxed{\text{Def } 9.22.}$ Let (X, d) be a metric space. if map: $\varphi: X \to X$. and theres a number $C < 1$. s.t. $d(\varphi(x), \varphi(y)) \leq C\, d(x, y)$. for all $x, y \in X$, then φ is said to be a Contraction of X into X.

$\boxed{\text{Thm } 9.23}$: If (X, d) is a Complete metric space. and if φ is a contraction then $\exists$ one and only one $x \in X$, s.t. $\varphi(x) = x$.

Pf: Pick arbitrary $x_0 \in X$, define a sequence $\{x_n\}_{n \geq 0} \subset X$, by $x_{n+1} = \varphi(x_n)$

Then by definition of Contraction we have: $d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq C\, d(x_n, x_{n-1})$

Simply by induction we have: $d(x_{n+1}, x_n) \leq C^n d(x_1, x_0)$, hence, for $m > n$.

$d(x_m, x_n) \leq \sum_{i=n+1}^{m} d(x_i, x_{i-1}) \leq (C^n + C^{n+1} + \cdots + C^{m-1})\, d(x_1, x_0) = \left(\sum_{i=n}^{m-1} C^i\right) d(x_1, x_0)$

Since $C < 1$, $\sum_{i=0}^{\infty} C^i$ converges. $\exists\, n_0$ s.t. $m > n \geq n_0$, $\Rightarrow \underline{\sum_{i=n}^{m-1} C^i < \frac{\varepsilon}{d(x_1, x_0)}}$ (Cauchy criterion)

Hence $d(x_m, x_n) < \frac{\varepsilon}{d(x_1, x_0)} \cdot d(x_1, x_0) = \varepsilon$. Since (X, d) is complete. , then $\{x_n\}$ is a convergent sequence. And φ is Contraction, then φ is uniformly continous.

$\Rightarrow \varphi(x) = \varphi(\lim_{n \to \infty} x_n) = \lim_{n \to \infty} \varphi(x_n) = \lim_{n \to \infty} x_{n+1} = x$. (we set $\lim_{n \to \infty} x_n = x$).

Thus we finish the proof. !

Proof for Inverse function theorem.

$\boxed{\text{Now we return to the Inverse function theorem:}}$

(i) Denote $A = \vec{f}'(\vec{a})$, by assumption A^{-1} exists (Invertible). put $\lambda = \frac{1}{2\|A^{-1}\|} \in (0, \infty)$

Since $\vec{f}'$ is Continuous at $\vec{a}$, ($\vec{f}' \in C'(E, \mathbb{R}^n)$.). $\exists\, \mu > 0$, s.t.

$|\vec{x} - \vec{a}| < \mu \Rightarrow \|\vec{f}'(\vec{x}) - \vec{f}'(\vec{a})\| = \|\vec{f}'(\vec{x}) - A\| < \lambda$ (By def of continuity).

Now we take $U = B_\mu(\vec{a})$. [Neigh... of $\vec{a}$. open ball with radius μ. , note that. open ball is a convex set.] Now for each $\vec{y} \in \mathbb{R}^n$ we consider a new function

$\Phi_{\vec{y}}: E \to \mathbb{R}^n$. $\Phi_{\vec{y}}(\vec{x}) = \vec{x} + A^{-1}(\vec{y} - \vec{f}(\vec{x}))$. $\Phi_{\vec{y}}(\vec{x}) \in C'(E, \mathbb{R}^n)$.

Hence, we have $\Phi'_{\vec{y}}(\vec{x}) = I_n + A^{-1}\vec{f}'(\vec{x}) = A^{-1}(A - \vec{f}'(\vec{x}))$. $I_n \in L(\mathbb{R}^n)$.

is the identity linear transformation. (Previous. we chain rule to show derivative of linear map)

Then $\forall \vec{x} \in E$. $\|\phi'_{\vec{y}}(\vec{x})\| = \|A^{-1}(A - \vec{f}'(\vec{x}))\| \underset{\text{by Thm 9.7}}{\leq} \|A^{-1}\| \cdot \|A - \vec{f}'(\vec{x})\|$

Now for $\vec{x} \in U$, (recall that $U = B_\mu(\vec{a})$, $\|\vec{f}'(\vec{x}) - A\| < \lambda$.) hence.

$\|\phi'_{\vec{y}}(\vec{x})\| \leq \|A^{-1}\| \cdot \|A - \vec{f}'(\vec{x})\| \leq \|A^{-1}\|\lambda = \frac{1}{2}$, $\underline{\|\phi'_{\vec{y}}(\vec{x})\| \leq \frac{1}{2}}$. since

U is a convex set. We shall use the MVT of several variables:

$\underline{\forall \vec{x}_1, \vec{x}_2 \in U, \ |\phi_{\vec{y}}(\vec{x}_1) - \phi_{\vec{y}}(\vec{x}_2)| \leq \frac{1}{2}|\vec{x}_1 - \vec{x}_2|}$, clearly U is complete. then $\phi_{\vec{y}}$

is a contraction, by Thm. 9.23 (contraction thm). $\exists$ one and only $\vec{x} \in U$. s.t.

$\phi_{\vec{y}}(\vec{x}) = \vec{x} \Rightarrow \vec{x} + A^{-1}(\vec{y} - \vec{f}(\vec{x})) = \vec{x} \Rightarrow A^{-1}(\vec{y} - \vec{f}(\vec{x})) = \vec{0} \Rightarrow \vec{y} - \vec{f}(\vec{x}) = \vec{0}$.

$\Rightarrow \vec{f}(\vec{x}) = \vec{y}$

$$\boxed{\text{Hence. we proved that } \forall \vec{y} \in \mathbb{R}^n. \text{ there's at most } \vec{x} \in U. \text{ with } \vec{f}(\vec{x}) = \vec{y}. \ \vec{f}_U \text{ is injective.}}$$

To finish the proof for (a). we still have to show: $\underline{V = \vec{f}(U). \text{ is open in } \mathbb{R}^n}$.

Let $\vec{y}_0 \in V$, we need to show, $\exists \rho > 0$. s.t. the open ball $B_\rho(\vec{y}_0) \subseteq V$

Now let $\vec{x}_0 \in U$. , $\vec{f}(\vec{x}_0) = \vec{y}_0$. this is unique by previous injection we obtained

Because U is open. $\exists r > 0$ s.t. the open ball $B_{2r}(\vec{x}_0) \subseteq U$. also. $\bar{B}_r(\vec{x}_0) \subseteq U$

So we have: $\vec{f}(\bar{B}_r(\vec{x}_0)) \subseteq V$. now we choose $\rho = \lambda r = \frac{r}{2\|A^{-1}\|}$

$\forall \vec{y} \in B_\rho(\vec{y}_0)$, $\exists \vec{x} \in \bar{B}_r(\vec{x}_0)$ with $\vec{f}(\vec{x}) = \vec{y}$. . let $\vec{y} \in B_\rho(\vec{y}_0)$ be arbitrary.

We have the associated: $\phi_{\vec{y}}: \bar{B}_r(\vec{x}_0) \to \mathbb{R}^n$, $\phi_{\vec{y}}(\vec{x}) = \vec{x} + A^{-1}(\vec{y} - \vec{f}(\vec{x}))$

To show $\exists \vec{x} \in \bar{B}_r(\vec{x}_0)$ with $\vec{f}(\vec{x}) = \vec{y}$ is equivalent to showing that $\phi_{\vec{y}}$

has a fixed point $\vec{x} \in \bar{B}_r(\vec{x}_0)$ $\Rightarrow$ this would follow directly from contraction principle!

If we show that $\phi_{\vec{y}}$ take $\bar{B}_r(\vec{x}_0)$ to itself., now the question reduced to:

$$\boxed{* \ \text{if } |\vec{x} - \vec{x}_0| \leq r. \text{ then } |\phi_{\vec{y}}(\vec{x}_0) - \vec{x}_0| \leq r.}$$

$$|\phi_{\vec{y}}(\vec{x}_0) - \vec{x}_0| = |A^{-1}(\vec{y} - \vec{y}_0)| \leq \|A^{-1}\| \, |\vec{y} - \vec{y}_0| < \|A^{-1}\| \cdot \rho \cdot = \frac{r}{2} \quad \cdots ①$$

Next. $|\Phi_{\vec y}(\vec x) - \Phi_{\vec y}(\vec y)| \le \frac{1}{2}|\vec x - \vec x_0| \le \frac{r}{2}$ ... ②

By ①, ②, (*) holds by triangle inequality.

Thus (1) part proof is finished.

$\boxed{\text{Now we continue the proof for part (2) for inverse function theorem}}$

Recall that we define $A = \vec f'(\vec a)$, and $\vec x \in U \Rightarrow \|\vec f'(\vec x) - A\| < \frac{1}{2\|A^{-1}\|} < \frac{1}{\|A^{-1}\|}$

From Thm 9.8. we know it is $\vec f'(\vec x) \in L(\mathbb{R}^n)$ is an invertible linear transformation.

To simplify, $T_{\vec x} = (\vec f'(\vec x))^{-1}$ $(\vec x \in U)$. [We want to show that the inverse $\vec g : V \to U$ of $\vec f$ is differentiable in V.]. If $\vec y \in V$ and $\vec g$ is diff at $\vec y$ and $\vec x = \vec g(\vec y)$

Now by chain rule: $\vec f'(\vec x)\,\vec g'(\vec y) = I_n \vec x = I_n$ (identity linear map)

then $\vec g'(\vec y) = (\vec f'(\vec x))^{-1} = T_{\vec x}$ (as we defined earlier).

We need to show: $\lim_{\vec k \to \vec 0} \frac{|\vec g(\vec x + \vec k) - \vec g(\vec y) - T_{\vec x}\vec k|}{|\vec k|} = 0$ ---- ③

So take $\vec k$, s.t. $\vec y + \vec k \in V$, and put $\vec h = \vec g(\vec y + \vec k) - \vec x$, so that:

$\vec x + \vec h = \vec g(\vec y + \vec k) \in U$, and $\vec f(\vec x + \vec h) = \vec y + \vec k$

Hence $\vec g(\vec x + \vec k) - \vec g(\vec y) - T_{\vec x}\vec k = \vec x + \vec h - \vec g(\vec y) - T_{\vec x}\vec k = \vec h - T_{\vec x}\vec k$

$= -T_{\vec x}\vec k + T_{\vec x}\cdot \vec f'(\vec x)\cdot \vec h = -T_{\vec x}(\underbrace{\vec f(\vec x + \vec h) - \vec f(\vec x)}_{\vec k} - \vec f'(\vec x)\cdot \vec h)$

Hence. (1) $|\vec g(\vec y + \vec k) - \vec g(\vec y) - T_{\vec x}\vec k| \le \|T_{\vec x}\| \cdot |\vec f(\vec x + \vec h) - \vec f(\vec x) - \vec f'(\vec x)\cdot \vec h|$

$\boxed{\text{More over: (2): } |\vec k| \ge \lambda |\vec h| = \frac{1}{2\|A^{-1}\|}\cdot|\vec h|}$. Still we use the $\Phi_{\vec y}$: (see prev def).

$\Phi_{\vec y}(\vec x + \vec h) - \Phi_{\vec y}(\vec x) = \vec x + \vec h + A^{-1}(\vec y - \vec f(\vec x + \vec h)) - \vec x - A^{-1}(\vec y - \vec f(\vec x))$

$= \vec h + A^{-1}(\vec f(\vec x) - \vec f(\vec x + \vec h)) = \vec h - A^{-1}\vec k$

$|\Phi_{\vec y}(\vec x + \vec h) - \Phi_{\vec y}(\vec x)| \le \frac{1}{2}|\vec h|$, $|\vec h| + |A^{-1}\vec k| < |\vec h - A^{-1}\vec k| < \frac{1}{2}|\vec h|$.

implies: $\frac{1}{2}|\vec h| \le |A^{-1}\vec k| \le \|A\|\,|\vec k|$, thus we proved (2).

From (1), (2) we get: $\dfrac{|\vec{g}(\vec{y}+\vec{k}) - \vec{g}(\vec{y}) - T\vec{k}|}{|\vec{k}|} \le \dfrac{\|T\| \, |\vec{f}(\vec{x}+\vec{h}) - \vec{f}(\vec{x}) - \vec{f}'(\vec{x})\vec{h}|}{\frac{1}{2\|A^{-1}\|} \cdot |\vec{h}|}$

$\lim_{\vec{h}\to\vec{0}}$ (right hand side) $=0$. by (2). $\lim_{\vec{k}\to\vec{0}} \vec{h} = \vec{0}$. Thus ③ is proved.

Now we proved that the g' is continous on V.

Invertibe liner transformation.

Note: g' is the composition of 3 mapping: $V \xrightarrow{\vec{g}} U \xrightarrow{\vec{f}'} \Omega \xrightarrow{\text{taking inverse}} \Omega \subset L(\mathbb{R}^n)$

$\Omega \subset L(\mathbb{R}^n)$.

Our next topic would be $\boxed{\text{Implicit function theorem}}$

Recall that in freshman calculus. if $x^2+y^2=1$. then $x(y) = \pm\sqrt{1-y^2}$

Next: Consider a general equation $f(x,y)=0$. , define a plane curve S:

let $(a,b)\in S$ be a point, $f(a,b)=0$. Assume: $\boxed{\text{i) } f\in C' \quad \text{ii) } \frac{\partial f}{\partial x}(a,b)\neq 0}$

Claim the following: 1) there exist an open U of (a,b). $U\subseteq E$.

2) An open interval $W\subset\mathbb{R}$, with $b\in W$. , 3) A function $g: W\to\mathbb{R}$. such that

$g(b)=a$. $(g(y), y)\in U$. , $\forall y\in W$. and $f(g(y),y)=0$, $\forall y\in W$.

(we write $x=g(y)$ and say that $f(x,y)=0$ " defines x as a function of y".)

Moreover. g is differentiable. with continous derivative: $\underline{g'(b) = -\dfrac{\frac{\partial f}{\partial y}(a,b)}{\frac{\partial f}{\partial x}(a,b)}}$

In fact, for all $y\in W$ sufficiently close to b, we have $\frac{\partial f}{\partial x}(g(y),y)\neq 0$ and

$$g'(y) = -\frac{\frac{\partial f}{\partial y}(g(y),y)}{\frac{\partial f}{\partial x}(g(y),y)} = -\frac{\frac{\partial f}{\partial y}(x,y)}{\frac{\partial f}{\partial y}(x,y)}.$$

Pf: Define $F: E\to\mathbb{R}^2$, $F(x,y) = (f(x,y), y) \Rightarrow F_1(x,y) = f(x,y)$. $F_2(x,y)=y$

Then $F\in C'(E,\mathbb{R}^2)$. $[F'(a,b)] = \begin{bmatrix} \frac{\partial f}{\partial x}(a,b) & \frac{\partial f}{\partial y}(a,b) \\ 0 & 1 \end{bmatrix}$, $\det[F'(a,b)] = \frac{\partial f}{\partial x}(a,b) \neq 0$

(by assumption).

$\boxed{\text{Now apply the Inverse theorem.}}$: $\exists$ open $U\subseteq E$, $(a,b)\in U$, $\exists$ open set $V\subseteq\mathbb{R}^2$.

$F(a,b)\in V$, s.t. $F: U\to V$ is a bijection.

Take $W \subset \mathbb{R}$. , $W = \{ y \in \mathbb{R} \mid (0,y) \in V \}$, i.e. $W = V \cap$ y-axis. , more precisely, denote by $j : \mathbb{R} \to \mathbb{R}^2$, $j(y) = (0,y)$. then $W = j^{-1}(V)$.

Since j is Continuous mapping, then W is open. and $b \in W$.

If $F^{-1} : V \to U$ is the inverse F, define : $g : W \to \mathbb{R}$ as the composition.

$pr_x \circ F^{-1} \circ j|_W$, where $pr_x : \mathbb{R}^2 - \mathbb{R}$, $pr_x(x,y) = x$. [projection onto x-axis]

Explicitly. for $y \in W$. $\Rightarrow (0,y) \in V$, then $F^{-1}(0,y) = (x,y)$ for unique x. and we get $g(y) = x$. Since $F^{-1} \in C'$ (by inverse function thm).

$\underbrace{g = pr_x \circ F^{-1}}$ $\circ j$ is of class C'

Use the Chain rule and Calculate $[F'(a,b)]^{-1}$ (inverse matrix).

We'll finish the proof. ($g'(b) = \cdots\cdots$)

$\boxed{\text{Remark:}}$ If we replace ii) with $\frac{\partial f}{\partial y}(a,b) \neq 0$. then the following holds:

$\exists$ open neighborhood, $(a,b) \in U$ and $a \in Z$, s.t. $\forall x \in Z$. there's a unique. $y \in \mathbb{R}$. with $(x,y) \in U$ and $f(x,y) = 0$, i.e. $\exists\, h : Z \to \mathbb{R}$. of class C' with $h(x) = y$. and $f(x, h(x)) = 0$, $h'(a) = -\dfrac{\frac{\partial f}{\partial x}(a,b)}{\frac{\partial f}{\partial y}(a,b)}$, Proof is the same.

Above is Case when $n=m=1$. of the implicit function theorem. , now "$n=1$, m=arbitrary"

$E \subseteq \mathbb{R}^{1+m}$ open., Consider : $f(x,\vec{y}) = f(x, y_1, \cdots y_m) = 0$. where $f : E \to \mathbb{R}$.

$(a, \vec{b}) = (a, b_1, \cdots b_m) \in E$. with $f(a,\vec{b}) = 0$

[Q : When is ~~the~~ the m-dimensional "surface" given by $f(x,\vec{y}) = 0$ the graph of a function $\boxed{x = g(\vec{y})}$ near $(a, \vec{b})$]

Again : i) $f \in$ class C' in E ii) $\frac{\partial f}{\partial x}(a, b_1, \cdots b_m) \neq 0$

Similarly : $\exists$ open sets $(a, \vec{b}) \in U \subset E \subset \mathbb{R}^{1+m}$. , $\vec{b} \in W \subset \mathbb{R}^m$.

Define : $g : W \to \mathbb{R}$. s.t. $(g(y_1, \cdots y_m), y_1, \cdots y_m) \in U$., $g(\vec{b}) = a$.

$f(g(y_1, \cdots y_m). y_1, y_2 \cdots y_m) = 0$. , also. $g \in C'$ and $\frac{\partial g}{\partial y_i}(b_1, \cdots b_m) = -\dfrac{\frac{\partial f}{\partial y_i}(a, b_1 \cdots b_m)}{\frac{\partial f}{\partial x}(a, b_1 \cdots b_m)}$

Pf: $F: E \to \mathbb{R}^{1+m}$. $F(x, \vec{y}) = (f(x, \vec{y}), \vec{y}) = (f(x, y_1$

$$[F'(a, \vec{b})] = \begin{bmatrix} \frac{\partial f}{\partial x}(a, \vec{b}), & \frac{\partial f}{\partial y_1}(a, \vec{b}) & \cdots & \frac{\partial f}{\partial y_m}(a, \vec{b}) \\ 0 & 1 & & \\ 0 & 0 & \ddots & \\ \vdots & 0 & 1 & \\ \vdots & \vdots & & \ddots \\ 0 & \cdot & 0 \quad 0 & 1 \end{bmatrix}$$

$\det [F'(a, b)] = \frac{\partial f}{\partial x}(a, \vec{b}) \neq 0$. Then we apply the Inverse function thm.

So $F: U \to V$ is bijection for appropriate U, V, $(a, \vec{b}) \in U$. $(0, \vec{b}) \in V$.

$j: \mathbb{R}^m \to \mathbb{R}^{1+m}$. : $j(y_1, \ldots y_m) = (0, y_1, \ldots, y_m)$

$pr_x: \mathbb{R}^{1+m} \to \mathbb{R}$: $pr_x(x, y_1, \ldots y_m) = x \Rightarrow$ projection on X-axis.

Take $W = j^{-1}(V) \subset \mathbb{R}^m$, put $g = pr_x \circ F^{-1} \circ j|_W$, then g satisfies the requirement.

$f(g(y_1, \ldots y_m), y_1 \ldots y_m) = 0$. differentiate this yields $\underline{g}$

Finally, we can extend to the general case: Implicit Function theorem.

$\boxed{\text{Theorem:}}$ Let $E \subseteq \mathbb{R}^{n+m}$ be open. $(\vec{a}, \vec{b}) \in E$. a point ($\vec{a} = (a_1, \cdots a_n) \in \mathbb{R}^n$, $\vec{b} = (b_1, \cdots b_m) \in \mathbb{R}^m$. Let $f: E \to \mathbb{R}^n$ be a function s.t. $f(\vec{a}, \vec{b}) = \vec{0} \in \mathbb{R}^n$.

i) $f \in C'(E, \mathbb{R}^n)$, ii) $A = f'(\vec{a}, \vec{b}) \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$. let $i: \mathbb{R}^n \to \mathbb{R}^{m+n}$

$i(\vec{h}) = (\vec{h}, \vec{0})$ be the inclusion. $i \in L(\mathbb{R}^n, \mathbb{R}^{m+n})$

$\underline{A_{\vec{x}} = A \circ i \in L(\mathbb{R}^n, \mathbb{R}^n)}$, we assume that $A_{\vec{x}}$ is an invertible linear transform.

$n \times n$ matrix:

$$[A_{\vec{x}}] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}, \vec{b}) & \frac{\partial f_1}{\partial x_2}(\vec{a}, \vec{b}) & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & & \vdots \\ & & \ddots & \\ \frac{\partial f_n}{\partial x_1}(\vec{a}, \vec{b}) & \cdots & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

is invertible by assumption.

Then there exists an open set $U \subset \mathbb{R}^{m+n}$. with $(\vec{a}, \vec{b}) \in U \subseteq E$. , there exists an open set $W \subset \mathbb{R}^m$. with $\vec{b} \in W$ and there exists $\vec{g}: W \to \mathbb{R}^n$, $\vec{g} \in C'$

$\vec{g}(\vec{b}) = \vec{a}$, $(\vec{g}(\vec{y}), \vec{y}) \in U$, $\forall \vec{y} \in W$, and $\vec{f}(\vec{g}(\vec{y}), \vec{y}) = \vec{0}$, $\forall \vec{y} \in W$.

and $\vec{f}(\vec{g}(\vec{y}), \vec{y}) = \vec{0}$. $[\vec{g}'(\vec{b})] = -(A_{\vec{x}})^{-1} A_{\vec{y}}$, where:

$A_{\vec{y}} \in L(\mathbb{R}^m, \mathbb{R}^n)$ is defined by

$$A_{\vec{y}} \vec{k} = A(\vec{0}, \vec{k})$$

(Note: $A_{\vec{y}} = \begin{bmatrix} \frac{\partial f_1}{\partial y_1}(\vec{a}, \vec{b}) & \cdots & \frac{\partial f_1}{\partial y_m}(\vec{a}, \vec{b}) \\ \vdots & & \\ \frac{\partial f_n}{\partial y_1}(\vec{a}, \vec{b}) & \cdots & \frac{\partial f_n}{\partial y_m}(\vec{a}, \vec{b}) \end{bmatrix}$)

pf: We shall treat it similar to the one "$n=1$". define: $F: E \to \mathbb{R}^{n+m}$.

$F(\vec{x}, \vec{y}) = (f(\vec{x}, \vec{y}), \vec{y})$, $F \in C'$ on E. and

$$[F'(\vec{a}, \vec{b})] = \begin{bmatrix} \underset{n\times n}{[A_{\vec{x}}]} & [A_{\vec{y}}] \\ \underset{m\times n}{[0]} & I_{m\times m} \end{bmatrix} . \quad \det[F'(\vec{a}, \vec{b})] = \det[A_{\vec{x}}] \neq 0$$

Now again apply the inverse function theorem. for appropriate open set.

$U, V \subset \mathbb{R}^{n+m}$. $F: U \to V$ is bijection. $F^{-1}: V \to U$ of class C'.

$U \ni (\vec{a}, \vec{b})$, $V \ni (\vec{0}, \vec{b})$

let $j: \mathbb{R}^m \to \mathbb{R}^{m+n}$, $j(\vec{y}) = (\vec{0}, \vec{y})$ be the inclusion, and $Pr_{\vec{x}}: \mathbb{R}^{n+m} \to \mathbb{R}^n$

$Pr_{\vec{x}}(\vec{x}, \vec{y}) = \vec{x}$ (projection). take $W = j^{-1}(V) \subset \mathbb{R}^m$. $g: W \to \mathbb{R}^n$

$\underline{\vec{g} = pr_{\vec{x}} \circ F^{-1} \circ j|_W}$, to formulate $\vec{g}'(\vec{b})$. differentiate:

$f(g(\vec{y}), \vec{y}) = \vec{0}$ using the chain rule.

Now let's move to Higher-order derivatives.

$\boxed{\text{Def } 9.39}$ Let $E \subset \mathbb{R}^n$ open, $f: E \to \mathbb{R}$. a function. Assume f has partial derivatives. $D_i f = \frac{\partial f}{\partial x_i}: E \to \mathbb{R}$. if each $D_j f$ is differentiable in E. we define the second partial derivative: $D_{ij} f = D_i D_j f = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial^2 f}{\partial x_i \partial x_j}: E \to \mathbb{R}$.

we can call this C^2 function.

$f \in C^k$ k^{th} differentiable. Continously.

Partial derivatives with different orders : Are they equal ? In general no! (Ex 27).

$\boxed{\text{Theorem 1. (Clairaut)}}$ If f is of class C^k, then all partial derivatives of orders $j \le k$ with respect to the same variables $x_{i_1}, \dots x_{i_j}$ are equal. $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$

$\boxed{\text{Theorem 2.}}$ Let $E \subset \mathbb{R}^2$ be open and let $f: E \to \mathbb{R}$, Suppose partial derivatives $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ exist at every point of E. let $(a,b) \in E$ be a point. If $\frac{\partial^2 f}{\partial y \partial x}$ is continuous at (a,b), then $\frac{\partial^2 f}{\partial x \partial y}$, exists at (a,b) and $\frac{\partial^2 f}{\partial y \partial x}(a,b) = \frac{\partial^2 f}{\partial x \partial y}(a,b)$

Pf: We start with 2 lemmae.

$\boxed{\text{Lemma 1.}}$ Let $h, k \in \mathbb{R}$, $h, k \neq 0$ s.t. the closed rectangle Q with vertices (a,b). $(a+h, b)$. $(a, b+k)$ and $(a+h, b+k)$., is contained in E.

put $\underline{\Delta(f, Q) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a,b)}$, Then there exist a point (x_0, y_0) in interior of Q. s.t. $\Delta(f, Q) = hk \cdot \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)$

Pf: Put $u: [a, a+h] \to \mathbb{R}$, $u(t) = f(t, b+k) - f(t, b)$

So by our previous definition. $\Delta(f, Q) = u(a+h) - u(a)$. Apply MVT

$\exists\ x_0 \in (a, a+h)$ with $\Delta(f, Q) = u(a+h) - u(a) = u'(x_0) \cdot h$ (by MVT)

Now $u'(x_0) = \frac{\partial f}{\partial x}(x_0, b+k) - \frac{\partial f}{\partial x}(x_0, b)$, take $v: [b, b+k] \to \mathbb{R}$, $\underline{v(s) = \frac{\partial f}{\partial x}(x_0, s)}$

Apply MVT to v again: $\exists\ y_0 \in (b, b+k)$. s.t. $v'(y_0) \cdot k = v(b+k) - v(b) = u'(x_0)$

So $\underline{\Delta(f, Q) = u'(x_0) \cdot h = (v'(y_0) k) h = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)}\, hk$, Continue to prove the theorem: Choose $\varepsilon > 0$, by continuity of $\frac{\partial^2 f}{\partial y \partial x}$ at (a,b). $\exists$ small open ball in E. $B \subset E$. , Centered at (a,b). s.t. $(x,y) \in B \Rightarrow \left| \frac{\partial^2 f}{\partial y \partial x}(x,y) - \frac{\partial^2 f}{\partial y \partial x}(a,b) \right| < \varepsilon$

Chose $h, k \neq 0$. , $\underline{\text{rectangle } Q \subset B. \ , (x_0, y_0) \in Q. \subset B}$, by the previous lemma we have: $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)\, hk = \Delta(f, Q)$. , $\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\Delta(f, Q)}{hk}$

Then $\left| \frac{\Delta(f, Q)}{hk} - \frac{\partial^2 f}{\partial y \partial x}(a,b) \right| < \varepsilon$. , ~~fixed~~ fix h, let $k \to 0$, $\lim_{k \to 0} \frac{\Delta(f, Q)}{k} = \frac{\partial f}{\partial y}(a+h, b) - \frac{\partial f}{\partial y}(a,b)$

Substitute: $\left| \frac{\frac{\partial f}{\partial y}(a+h, b) - \frac{\partial f}{\partial y}(a,b)}{h} - \frac{\partial^2 f}{\partial y \partial x}(a,b) \right| < \varepsilon.$

$\uparrow$

This guy is by definition: $\frac{\partial f}{\partial x \partial y}(a,b)$.

Now: Differentiation of integrals:

Question: Under what Conditions: $\frac{d}{dt}\int_a^b \varphi(x,t)\,dt = \int_a^b \frac{\partial \varphi}{\partial t}(x,t)\,dx$?

[Ex28. gives ~~that~~ equation might not hold]

$\boxed{\text{Theorem 9.42}}$ From rudin directly.

9.42 Theorem *Suppose*

(a) $\varphi(x, t)$ *is defined for* $a \le x \le b$, $c \le t \le d$;
(b) α *is an increasing function on* $[a, b]$;

For each $t \in [c,d]$. denote $\varphi_t : [a,b] \to \mathbb{R}$

$\varphi_t(x) = \varphi(x,t)$. Also fix a point $s \in (c,d)$

(c) $\varphi^t \in \mathscr{R}(\alpha)$ *for every* $t \in [c, d]$;
(d) $c < s < d$, *and to every* $\varepsilon > 0$ *corresponds a* $\delta > 0$ *such that*

$$|(D_2 \varphi)(x, t) - (D_2 \varphi)(x, s)| < \varepsilon$$

for all $x \in [a, b]$ *and for all* $t \in (s - \delta, s + \delta)$.

Assume the following:

1) $\varphi_t \in \mathscr{R}(\alpha)$. $\forall t \in [c,d]$

2) $\left| \frac{\partial \varphi}{\partial t}(x,t) - \frac{\partial \varphi}{\partial t}(x,s) \right| < \varepsilon$.

Define

(100) $$\underline{f(t) = \int_a^b \varphi(x, t)\, d\alpha(x)} \qquad (c \le t \le d).$$

Then $(D_2 \varphi)^s \in \mathscr{R}(\alpha)$, $f'(s)$ *exists, and*

(101) $$f'(s) = \int_a^b (D_2 \varphi)(x, s)\, d\alpha(x).$$

i) $\frac{\partial \varphi}{\partial t}(x,s) \in \mathscr{R}(\alpha)$ ii) $f'(s)$ exists

iii) $f'(s) = \int_a^b \frac{\partial \varphi}{\partial t}(x,s)\,d\alpha(x)$.

Note that (c) simply asserts the existence of the integrals (100) for all $t \in [c, d]$. Note also that (d) certainly holds whenever $D_2 \varphi$ is continuous on the rectangle on which φ is defined.

Pf: Let $t_n \to s$, pick N. s.t. $n \ge N$. $\Rightarrow |t_n - s| < \delta$

$\varphi^x(t) = \varphi(x,t)$. $\Rightarrow \exists\ u_{n,x}$ with $\frac{\partial \varphi}{\partial t}(x, u_{n,x}) = \frac{\varphi(x,t_n) - \varphi(x,s)}{t_n - s}$.

Since $|u_{n,x} - s| < |t_n - s| < \delta$. $\Rightarrow \left| \frac{\partial \varphi}{\partial t}(x, u_{n,x}) - \frac{\partial \varphi}{\partial t}(x,s) \right| < \varepsilon$ (by (d))

Hence $\left| \frac{\varphi(x,t_n) - \varphi(x,s)}{t_n - s} - \frac{\partial \varphi}{\partial t}(x,s) \right| < \varepsilon$ $\forall x \in [a,b]$, $\forall n \ge N_\varepsilon$

i.e. $\underline{\lim_{n\to\infty} \frac{\varphi(x,t_n) - \varphi(x,s)}{t_n - s}} = \frac{\partial \varphi}{\partial t}(t,s)$ $\to$ uniformly.

Each $\frac{\varphi(x,t_n) - \varphi(x,s)}{t_n - s} \in \mathscr{R}(\alpha)$ on $[a,b]$. . hence by uniform convergence.

$\frac{\partial \varphi}{\partial t}(x,s) \in \mathscr{R}(\alpha)$. on $[a,b]$ $\therefore$ i).

and $\lim_{n\to\infty} \int_a^b \frac{\varphi_{t_n} - \varphi_s}{t_n - s} \, d\alpha(x) = \int_a^b \frac{\partial \varphi}{\partial t}(x,s) \, d\alpha(x)$ ... ①

By definition: $\frac{f(t_n) - f(s)}{t_n - s} = \int_a^b \frac{\varphi_{t_n} - \varphi_s}{t_n - s} \, d\alpha(x)$.

① says that f is differentiable at s and $f'(s) = \lim_{n\to\infty} \frac{f(t_n) - f(s)}{t_n - s} =$ ①

$$= \int_a^b \frac{\partial \varphi}{\partial t}(x,s) \, d\alpha(x).$$

(i), (ii) proof finished!