Mapping Theorems and the Implicit Function Theorem

- $\langle Ax, y\rangle = (Ax)^T y = \sum_{r=1}^{m} (row_r \cdot x)\, y_r = \sum_{r=1}^{m}\sum_{c=1}^{n} A_{rc} x_c y_r$
- $h(x) = \langle f(x), z\rangle = z^T f(x)$, then $dh(x) = z^T df(x)$.

Matrix and Linear Map

- Let A be any $m \times n$ matrix, and $|A| = \sqrt{\sum_{c=1}^{n}\sum_{r=1}^{m} |A_{rc}|^2}$
  - $\forall x \in \mathbb{R}^n \; |Ax| \le |A||x|$

    Proof. $\left|A\sum_{j=1}^{n} x_j e^{(j)}\right| = \left|\sum_{j=1}^{n} x_j A e^{(j)}\right| \le \sum_{j=1}^{n} x_j \left|Ae^{(j)}\right| \overset{C.S.}{\le} \sqrt{\sum_{j=1}^{n} x_j^2}\sqrt{\sum_{j=1}^{n}\left|Ae^{(j)}\right|^2} = |x||A|$.
  - $\forall x \in \mathbb{R}^n \; Ax = 0 \;\Leftrightarrow\; A = 0$ matrix.

    Proof. "$\Rightarrow$" Choose $x = e^{(j)}$. Then, Ax is the j^{th} column of A. $Ax = 0$ implies the j^{th} column of A is zero.
  - Let $\|A\| = \min\{c : \forall x \in \mathbb{R}^n \; |Ax| \le c|x|\}$.
    - $|A_{jk}| \le \|A\| \le |A|$.
- Let A be any $n \times n$ matrix
  - $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ is continuous.
  - A^{-1}, if exists, is a continuous function of the entries of A.

    Proof. Cramer's rule.
- **Linear map**: $L_A(x) = Ax : \mathbb{R}^n \to \mathbb{R}^m$, where A is an $m\times n$ matrix.
  - $n = \dim(\mathbb{R}^n) \ge \dim(\{Ax : x \in \mathbb{R}^n\})$

- L_A or $A_{m\times n}$ is **injective**

$\equiv$ $\forall x \in \mathbb{R}^n$, $L_A(x_1) = L_A(x_2) \Leftrightarrow x_1 = x_2$.

$\equiv$ $\exists r > 0 \; \forall x \in \mathbb{R}^n \; r|x| \le |L(x)|. \Rightarrow (r|x| \le |Ax| \le |A||x|)$

Proof. "$\Leftarrow$" $Ax_1 = Ax_2 \Rightarrow A(x_1 - x_2) = 0$. So, $0 = |A(x_1 - x_2)| \ge r|x_1 - x_2|$. Thus, $|x_1 - x_2| = 0$. "$\Leftarrow$" Because $\|x\| = |L(x)|$ is a norm on $\mathbb{R}^n$, it is equivalent to $|x|$.

$\equiv$ $L(x) = 0$ iff $x = 0$.

Proof. "$\Rightarrow$" Because $A0 = 0$, so, $Ax = 0 = A0 \Rightarrow x = 0$. "$\Leftarrow$" $Ax_1 = Ax_2 \Rightarrow$ $A(x_1 - x_2) = 0 \Rightarrow \; x_1 - x_2 = 0$.

≡ $Ax = 0$ has only the trivial solution.

≡ $\forall b\ \ Ax = b$ has at most one solution.

≡ The n columns of A are linearly independents. $\Rightarrow n \le m$ $\boxed{1:1}$

$\Rightarrow\ \|x\| = |L(x)| = |Ax|$ is a norm on $\mathbb{R}^n$.

Proof. 1) $|\cdot| \ge 0$. $x = 0 \Rightarrow Ax = 0$. Because L is injective, $Ax = 0 \Rightarrow x = 0$. 2) $|A(ax)| = |a||Ax|$. 3) $|A(x+y)| \le |Ax| + |Ay|$.

- Def: L_A or $A_{m\times n}$ is **surjective** if $\forall y \in \mathbb{R}^m\ \exists x \in \mathbb{R}^n$ such that $L_A(x) = y$.

≡ $\forall b \in \mathbb{R}^m$ the equation $Ax = b$ always has at least one solution.

≡ The n columns of A span $\mathbb{R}^m \Rightarrow n \ge m$. $\boxed{\text{onto}}$

- Def: L_A or $A_{m\times n}$ is **bijective** if it is both injective and surjective. (1:1 and onto).

≡ A is invertible.

≡ A is surjective ≡ A is injective.

Mean Value theorem

- Def: Let (a,b) denote the line segment joining a and b.

- MVT0: Mean Value theorem:

  Let open $\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$ differentiable.

  $[a,b] \subset \Omega \Rightarrow \exists c \in (a,b)$ such that $f(b) - f(a) = df(c)(b-a) = \nabla f(c)\cdot(b-a)$.

- MVT1: Let open $\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}^m$ differentiable.

  $[a,b] \subset \Omega \Rightarrow \exists c \in (a,b)$ such that $|f(b) - f(a)| \le \left|\underset{m\times n}{df(c)}(b-a)\right|$.

Proof. Let $y = f(b) - f(a)$, fixed. If $y = 0$, then done because $|df(c)(b-a)| \ge 0$. For $y \ne 0$, consider $h(x) = \left\langle f(x), \frac{y}{|y|}\right\rangle = \frac{1}{|y|}\sum_{k=1}^{m} f_k(x) y_k$. Then $dh(c) = \left[\frac{\partial h}{\partial x_1}(c) \cdots \frac{\partial h}{\partial x_n}(c)\right]$ with $\frac{\partial h}{\partial x_j}(x) = \frac{1}{|y|}\sum_{k=1}^{m} y_k \frac{\partial f_k}{\partial x_j}(x) = \frac{1}{|y|}\sum_{k=1}^{m} y_k [df(x)]_{kj}$. So,

$$dh(c)(b-a) = \sum_{j=1}^{n} \frac{\partial h}{\partial x_j}(c)(b-a)_j = \sum_{j=1}^{n} \frac{1}{|y|}\sum_{k=1}^{m} y_k [df(c)]_{kj} (b-a)_j$$

$$= \frac{1}{|y|}\langle df(c)(b-a), y\rangle \le \frac{1}{|y|}|df(c)(b-a)||y| \text{ by C.S.}$$

- MVT2: Let open $\Omega \subset \mathbb{R}^n$, $f:\Omega \to \mathbb{R}^m$ differentiable.
  If $[a,b] \subset \Omega$, $x_0 \in \Omega$, then $\left| f(b) - f(a) - df(x_0)(b-a) \right| \le |b-a| \sup_{c \in (a,b)} \left| df(c) - df(x_0) \right|$.

Proof. Let $g(x) = f(x) - df(x_0)x : \Omega \to \mathbb{R}^m$. Then, $dg(x) = df(x) - df(x_0)$. From above, $\exists c \in (a,b)$ $\left| g(b) - g(a) \right| \le \left| dg(c)(b-a) \right|$. $\left| g(b) - g(a) \right| = f(b) - f(a)$.

$\left| dg(c)(b-a) \right| = \left| (df(c) - df(x_0))(b-a) \right| \le \left| df(c) - df(x_0) \right| |b-a|$

$\le |b-a| \sup_{c \in (a,b)} \left| df(c) - df(x_0) \right|$.

- MVT3: Let open $\Omega \subset \mathbb{R}^n$, $f:\Omega \to \mathbb{R}^m$ is C^1. $x_0 \in \Omega$
  $\forall e > 0\ \exists d_e > 0$ such that $\forall x_1 \forall x_2\ x_1, x_2 \in B_{d_e}(x) \Rightarrow$ 1) $x_1, x_2 \in \Omega$, and 2)
  $\left| f(x_2) - f(x_1) - df(x_0)(x_2 - x_1) \right| \le e |x_2 - x_1|$.

- MVT3′: Let open $\Omega \subset \mathbb{R}^n$, $f:\Omega \to \mathbb{R}^m$ is C^1. $x_0 \in \Omega$
  $\forall e > 0\ \exists d_e > 0$ such that $\forall x_1 \forall x_2\ |x_k - x_0|$ £ d_e, $k = 1,2 \Rightarrow$ 1) $x_1, x_2 \in \Omega$, and 2)
  $\left| f(x_2) - f(x_1) - df(x_0)(x_2 - x_1) \right| \le e |x_2 - x_1|$.

Proof. Given $e > 0$. Ω open implies $\exists d_1 > 0\ \forall d\ \ 0 < d < d_1 \Rightarrow B_d(x_0) \subset \Omega$. By continuity of $df(x)$. Given $e > 0$, $\exists d_2 > 0\ \forall c \in B_{d_2}(x_0)$ $\left| df(c) - df(x_0) \right| \le e$.

Choose $d = \min(d_1, d_2)$, then $\forall x_1 \forall x_2\ \ x_1, x_2 \in B_d(x) \Rightarrow$ 1) $x_1, x_2 \in \Omega$ and $(x_1, x_2) \subset \Omega$ because $0 < d < d_1$, and open ball is convex, 2) because from above $\left| f(x_2) - f(x_1) - df(x_0)(x_2 - x_1) \right| \le |x_2 - x_1| \sup_{c \in (x_1, x_2)} \left| df(c) - df(x_0) \right|$, and $d < d_2$ implies that an upperbound of $\left| df(c) - df(x_0) \right|$ is e.

Mapping Theorems

- **The injective mapping theorem**: Let open $\Omega \subset \mathbb{R}^n$, $f:\Omega \to \mathbb{R}^m$ is C^1, $c \in \Omega$. $df(c)$ is injective $\Rightarrow$ $\exists d > 0$ such that $f|_{B_d(c)}$ is injective.

Proof. $df(c)$ is injective; thus, $\exists r > 0\ \forall x \in \mathbb{R}^n\ \ r|x| \le |df(c)x|$. Note that from triangle inequality $\left| f(x_2) - f(x_1) - df(c)(x_2 - x_1) \right| \ge \left| df(c)(x_2 - x_1) \right| - \left| f(x_2) - f(x_1) \right|$

$\ge r|x_2 - x_1| - \left| f(x_2) - f(x_1) \right|$. From MVT3, choose $e = \frac{r}{2}$ and $x_0 = c$. $\exists d > 0$ such that

$\forall x_1 \forall x_2\ \ x_1, x_2 \in B_d(c) \Rightarrow \frac{r}{2}|x_2 - x_1| \ge \left| f(x_2) - f(x_1) - df(x_0)(x_2 - x_1) \right|$ which is

$\geq r|x_2 - x_1| - |f(x_2) - f(x_1)|$. So, $|f(x_2) - f(x_1)| \geq \frac{r}{2}|x_2 - x_1|$. Thus, $f(x_2) = f(x_1) \Rightarrow$ $\frac{r}{2}|x_2 - x_1| = 0$. $f(x_2), f(x_1)$ are well defined because $x_1, x_2 \in \Omega$.

- $\exists r > 0 \;\; \forall x_1 \, \forall x_2 \;\; x_1, x_2 \in B_d(c) \Rightarrow |f(x_2) - f(x_1)| \geq \frac{r}{2}|x_2 - x_1|$.
- $f|_{B_d(c)} : B_d(c) \xrightarrow[\text{onto}]{1:1} f(B_d(c))$.

  Thus, well defined inverse $g = \left(f|_{B_d(c)}\right)^{-1} : f(B_d(c)) \xrightarrow[\text{onto}]{1:1} B_d(c)$.
- $g = \left(f|_{B_d(c)}\right)^{-1} : f(B_d(c)) \xrightarrow[\text{onto}]{1:1} B_d(c)$ is uniformly continuous.

  Proof. Let $y_1, y_2 \in f(B_d(c))$, then $\exists x_1, x_2 \in B_d(c)$ such that $y_1 = f(x_1)$ and $y_2 = f(x_2)$ $(x_1 = g(y_1)$ and $x_2 = g(y_2))$. $|f(x_2) - f(x_1)| \geq \frac{r}{2}|x_2 - x_1| \Rightarrow$ $|g(y_2) - g(y_1)| \leq \frac{2}{r}|y_2 - y_1|$.

$\Rightarrow$ $\exists d_1 > 0$ such that $B_{d_1}(c) \subset \Omega$ and $\forall x \in B_{d_1}(c)$ $df(x)$ is injective.

Proof. $df(c)$ is injective; so, $\exists r > 0 \; r|y| \leq |df(c)y|$. By continuity of $df(x)$, $\exists d_1 > 0 \;\; |x - c| < d_1 \Rightarrow |df(c) - df(x)| \leq \frac{r}{2}$. Hence, $|df(x)y|$ $\geq |df(c)y| - |df(c)y - df(x)y| \geq r|y| - |df(c) - df(x)||y| \geq r|y| - \frac{r}{2}|y|$.

$\Rightarrow$ $\exists d_2 > 0$ such that

1) $f|_{B_{d_2}(c)}$ is injective
2) $\forall x \in B_{d_2}(c)$ $df(x)$ is injective.
3) $g = \left(f|_{B_d(c)}\right)^{-1} : f(B_d(c)) \xrightarrow[\text{onto}]{1:1} B_d(c)$ is unique, and uniformly continuous.
4) $\exists r > 0 \;\; \forall x_1 \, \forall x_2 \;\; x_1, x_2 \in B_d(c) \Rightarrow |f(x_2) - f(x_1)| \geq \frac{r}{2}|x_2 - x_1|$.

Proof. Take $d_2 = \min(d, d_1) > 0$.

- $\underset{m\times n}{df(c)}$ is surjective. ($L_{df(c)}(x) = df(c)x : \mathbb{R}^n \to \mathbb{R}^m$ is surjective)

$\equiv$ (Def) $\forall y \in \mathbb{R}^m \; \exists x \in \mathbb{R}^n$ such that $L_{df(c)}(x) = df(c)x = y$.

$\Rightarrow$ $\exists \underset{n\times m}{M} \;\; \underset{m\times n}{df(c)} \underset{n\times m}{M} = \underset{m\times m}{I}$, i.e., $L_{df(c)} \circ L_M(x) = x$. (identity mapping)

Proof. Let $e^{(1)},\ldots,e^{(m)}$ be the canonical basis in $\mathbb{R}^m$, then $\forall i\in\{1,\ldots,m\}$ $\exists u^{(i)}\in\mathbb{R}^n$ such that $e^{(i)}=df(c)u^{(i)}$. Also, $\forall x\in\mathbb{R}^m$ $x=\sum_{i=1}^m x_i e^{(i)}$. Define $L_M(x)=\sum_{i=1}^m x_i u^{(i)}:\mathbb{R}^m\to\mathbb{R}^n$. Then, $\underset{n\times m}{M}=\left[u^{(1)}\cdots u^{(m)}\right]$. And

$$L_{df(c)}\circ L_M(x)=L_{df(c)}(L_M(x))=df(c)Mx=df(c)\left(\sum_{i=1}^m x_i u^{(i)}\right)=\sum_{i=1}^m x_i df(c)u^{(i)}$$

$$=\sum_{i=1}^m x_i df(c)u^{(i)}=\sum_{i=1}^m x_i e^{(i)}=x.$$

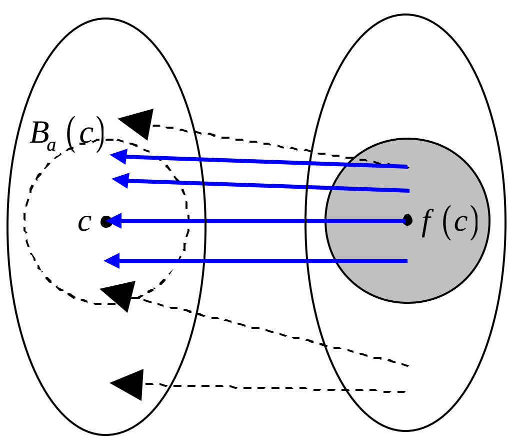
- Note also that $|Mx|=\left|\sum_{i=1}^m x_i u^{(i)}\right|\le\sum_{i=1}^m |x_i||u^{(i)}|\underset{C.S.}{\le}\left(\sqrt{\sum_{j=1}^m |u^{(i)}|^2}\right)|x|=b|x|$

- The **surjective mapping theorem**: Let open $\Omega\subset\mathbb{R}^n$, $f:\Omega\to\mathbb{R}^m$ is C^1, $c\in\Omega$. $df(c)$ is surjective
$\Rightarrow\exists e_1,e_2>0$ such that $B_{e_1}(c)\subset\Omega$ and $B_{e_2}(f(c))\subset f(B_{e_1}(c))$.

$\Rightarrow\exists a,b>0$ such that $\forall y\in\mathbb{R}^m$ $|y-f(c)|\le\frac{a}{2b}$, $\exists x\in\mathbb{R}^n$ $|x-c|\le a$ such that $f(x)=y$.

$\Rightarrow\exists a,b>0$ such that $B_{\frac{a}{2b}}(f(c))\subset f(B_a(c))$.

$\Rightarrow\exists a,b>0$ such that $\forall y\in\mathbb{R}^m$ $|y-f(c)|<\frac{a}{2b}$, $\exists x\in\mathbb{R}^n$ $|x-c|<a$ such that $f(x)=y$.



Proof. From MVT3′, let $e=\frac{1}{2b}$, $x_0=c$. Take $a=d$. Because $df(c)$ is surjective, $\exists M$ $df(c)M=I$. Also, $\exists b>0$ $|Mx|\le b|x|$. Given y, $|y-f(c)|\le\frac{a}{2b}$, will construct sequence $x^{(n)}\to x$, $|x-c|\le a$ and $f(x)=y$. Let $x_0=c$.

Define

$x^{(1)} = x^{(0)} - M\left(f\left(c\right) - y\right)$.

For $\ell \geq 1$, $x^{\ell+1} = x^{\ell} - M\left(f\left(x^{\ell}\right) - f\left(x^{\ell-1}\right) - df\left(c\right)\left(x^{\ell} - x^{\ell-1}\right)\right)$.

Claim: $\forall k \in \mathbb{N}$ 1) $\left|x^{(k)} - x^{(k-1)}\right| \leq \frac{a}{2^k}$, and 2) $\left|x^{(k)} - c\right| \leq \left(1 - \frac{1}{2^k}\right)a < a$

Proof. For k = 1, $\left|x^{(1)} - x^{(0)}\right| = \left|M\left(f\left(c\right) - y\right)\right| \leq b\,\frac{a}{2b} = \frac{a}{2^1}$. Also,

$\left|x^{(1)} - c\right| = \left|x^{(1)} - x^{(0)}\right| \leq \frac{a}{2^1} = \left(1 - \frac{1}{2^1}\right)a$.

Assume true for k = 1, .., ℓ. Then by 2), $x_{\ell}, x_{\ell-1} \in B_d\left(c\right)$, thus

$\left|f\left(x^{\ell}\right) - f\left(x^{\ell-1}\right) - df\left(c\right)\left(x^{\ell} - x^{\ell-1}\right)\right| \leq \frac{1}{2b}\left|x^{\ell} - x^{\ell-1}\right|$. Hence,

$\left|x^{\ell+1} - x^{\ell}\right| = \left|-M\left(f\left(x^{\ell}\right) - f\left(x^{\ell-1}\right) - df\left(c\right)\left(x^{\ell} - x^{\ell-1}\right)\right)\right| \leq b\,\frac{1}{2b}\left|x^{\ell} - x^{\ell-1}\right|$

$\leq \frac{1}{2}\frac{a}{2^{\ell}} = \frac{a}{2^{\ell+1}}$. Also, $\left|x^{(\ell+1)} - c\right| < \left|x^{(\ell+1)} - x^{(\ell)}\right| + \left|x^{(\ell)} - c\right| \leq \frac{a}{2^{\ell+1}} + \left(1 - \frac{1}{2^{\ell}}\right)a$

$= \left(1 - \frac{1}{2^{\ell+1}}\right)a$ ..

Thus, $x^{(n)} \in B_d\left(c\right)$. Also, $\left(x^{(n)}\right)$ is Cauchy. Assume $k > \ell$, then by 1)

$\left|x^{\ell} - x^{k}\right| = \left|\sum_{i=\ell}^{k-1}\left(x_i - x_{i+1}\right)\right| \leq \sum_{i=\ell}^{k-1}\left|x_i - x_{i+1}\right| \leq \sum_{i=\ell}^{k-1}\frac{a}{2^{i+1}} < \sum_{i=\ell}^{\infty}\frac{a}{2^{i+1}} = \frac{a}{2^{\ell}}$. Also, by 2),

$\left|x - c\right| = \lim_{n\to\infty}\left|x^{(n)} - c\right| \leq \lim_{n\to\infty}\left(1 - \frac{1}{2^n}\right)a = a$.

Claim: $df\left(c\right)\left(x^{\ell+1} - x^{\ell}\right) = y - f\left(x_{\ell}\right)$.

$df\left(c\right)\left(x^{(1)} - x^{(0)}\right) = y - f\left(x_0\right)$. By induction,

$$df\left(c\right)\left(x^{\ell+1} - x^{\ell}\right) = -f\left(x^{\ell}\right) + f\left(x^{\ell-1}\right) + df\left(c\right)\left(x^{\ell} - x^{\ell-1}\right)$$

$$= -f\left(x^{\ell}\right) + f\left(x^{\ell-1}\right) + y - f\left(x_{\ell-1}\right) = y - f\left(x_{\ell}\right)$$

Thus, $0 \leq \left|y - f\left(x\right)\right| = \lim_{\ell\to\infty}\left|y - f\left(x_{\ell}\right)\right| = \lim_{\ell\to\infty}\left|df\left(c\right)\left(x^{\ell+1} - x^{\ell}\right)\right| \leq \left|df\left(c\right)\right|0 = 0$.

Proof. Take $a' > a$. Also, take b' large enough such that $\frac{a'}{2b'} < \frac{a}{b}$. Then

$\forall y \in \mathbb{R}^m$ $\left|y - f\left(c\right)\right| < \frac{a'}{2b'} \leq \frac{a}{2b}$ $\left|y - f\left(c\right)\right| \leq \frac{a}{2b}$ $\exists x \in \mathbb{R}^n$ $\left|x - c\right| \leq a < a'$ such that

$f\left(x\right) = y$.

- **Open mapping theorem** Let open $\Omega \subset \mathbb{R}^n$. $f:\Omega \to \mathbb{R}^m$ is C^1. $\forall x \in \Omega$ $df(x)$ is surjective. Then, f is an open mapping. $\left(\forall \text{ open } G \subset \Omega,\ f(G) \text{ is open in } \mathbb{R}^m\right)$

Proof. Let $b \in f(G)$. Then, $\exists c \in G$ $f(c) = b$. Consider f on open G. $df(c)$ is surjective. By the SMT, $\exists e_1, e_2 > 0$ such that $B_{e_1}(c) \subset G$ and $B_{e_2}(f(c)) \subset f\left(B_{e_1}(c)\right) \subset f(G)$.

- **Inversion (Mapping) Theorem**: Let open $\Omega \subset \mathbb{R}^n$, $f:\Omega \to \mathbb{R}^n$ is C^1, $c \in \Omega$. $\underset{n\times n}{df(c)}$ is bijective, then $\exists U$ open neighborhood of c such that
  1) $V = f(U)$ is an open neighborhood of $f(c)$.
  2) $f|_U : U \to V$ is bijective.
  3) $g = \left(f|_U\right)^{-1} : V \to U$ is C^1.
  4) $\forall y \in V$ $dg(y) = \left[df(g(y))\right]^{-1}$.

Proof 2): By injective mapping theorem, because $df(c)$ is injective, $\exists U = B_d(c)$ such that $f|_U$ is injective and $\forall x \in U$ $df(x)$ is injective ($\Rightarrow$ bijective $\Rightarrow$ surjective).
Also, unique $g = \left(f|_U\right)^{-1} : V \xrightarrow[\text{onto}]{1:1} U$ is continuous; $V = f(U)$.
Proof 1): Because $\forall x \in U$ $df(x)$ is surjective, by open mapping theorem, $V = f(U)$ is open. $c \in U \Rightarrow f(c) \in V$.
Proof 4): Let $y, y_0 \in V$. Then, $\exists$ unique $x, x_0 \in U$, $f(x) = y$, $g(y) = x$, $f(x_0) = y_0$, $g(y_0) = x_0$. $x_0 \in U \Rightarrow \left[df(x_0)\right]^{-1}$ exists. f is differentiable at x_0, thus,

$$f(x) - f(x_0) = df(x_0)(x - x_0) + o(|x - x_0|) \text{ as } x \to x_0.$$

$\times\left[df(x_0)\right]^{-1}$, and get

$$\left[df(x_0)\right]^{-1}(y - y_0) = g(y) - g(y_0) + \left[df(x_0)\right]^{-1} o(|x - x_0|).$$

Claim: $r(x) = o(|x - x_0|)$ $x \to x_0$, then $Ar(x) = o(|y - y_0|)$ as $y \to y_0$.

By continuity of g, as $y \to y_0$, we have $x \to x_0$. Note that $|Ar(x)| \le |A||r(x)|$, and by the IMT, $|y - y_0| = |f(x) - f(x_0)| \ge \frac{r}{2}|x - x_0|$. Thus, $\frac{|Ar(x)|}{|y - y_0|} \le \frac{|A||r(x)|}{\frac{r}{2}|x - x_0|}$.

And $\lim_{x\to x_0} \frac{|r(x)|}{|x - x_0|}$, we have $\lim_{y\to y_0} \frac{|Ar(x)|}{|y - y_0|} = \lim_{x\to x_0} \frac{|A||r(x)|}{\frac{r}{2}|x - x_0|} = 0$.

Proof 3): $dg(y) = [df(g(y))]^{-1}$ is C^1. By inversion theorem, $g(y)$ is continuous. df is continuous. So, $df(g(y))$ is continuous. Inverse is also continuous (already know that determinant $\neq 0$).

- $f \circ g(y) = x : V \to V$, $g \circ f(x) = y : U \to U$.

- Inverse Function Theorem: Let $f : (a,b) \to \mathbb{R}$ be C^1, $f((a,b)) = (c,d)$; and suppose either $f'(x) > 0$ or $f'(x) < 0$ on (a,b). Then $f^{-1} : (c,d) \xrightarrow[\text{onto}]{} (a,b)$ is C^1, and $\left(f^{-1}\right)'(y) = \dfrac{1}{f'(x)}$ if $y = f(x)$.

Implicit Function Theorem

- **Implicit Function Theorem**: Let $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ open. $F : \Omega \to \mathbb{R}^m$ C^1. $(x_0, y_0) \in \Omega$, $F(x_0, y_0) = 0$, $d_y F(x_0, y_0)$ is invertible, then
  1) $\exists W$ open neighborhood of x_0, and a unique function f of class C^1, $f : W \to \mathbb{R}^m$, $f(x_0) = y_0$, and $\forall x \in W$ $F(x, f(x)) = 0$.
  2) Also, $\exists U$ open neighborhood of (x_0, y_0) such that if $(x,y) \in U$ and $F(x,y) = 0$ then $x \in W$ and $y = f(x)$.
  3) Moreover, $\exists r > 0$ $|x - x_0| < r \Rightarrow df(x) = -\left(d_y F(x, f(x))\right)^{-1} d_x F(x, f(x))$.

- We want to consider a system of m equations for m unknown functions $y_1, y_2, \ldots, y_m$ (each being a function of n variables $x_1, \ldots, x_n$.)
- Notice that we are prejudging the outcome that the number of equations and unknowns should be equal if there is to be any hope of having unique solutions.
- Replace F be its best affine approximation at (x_0, y_0),

$$F(x,y) = \overset{0}{\cancel{F(x_0, y_0)}} + dF(x_0, y_0)((x,y) - (x_0, y_0))$$
$$= d_x F(x_0, y_0)(x - x_0) + d_y F(x_0, y_0)(y - y_0)$$

$$F(x,y) = 0 \Rightarrow$$

$$\left(d_y F(x_0, y_0)\right) y = \boxed{Ay = b(x)} = d_x F(x_0, y_0)(x - x_0) + d_y F(x_0, y_0)\, y_0$$

- $\underset{m\times(n+m)}{dF(x,y)} = \begin{bmatrix} d_x F(x,y) & d_y F(x,y) \end{bmatrix}$

$$= \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(x,y) & \cdots & \frac{\partial F_1}{\partial x_n}(x,y) & \frac{\partial F_1}{\partial y_1}(x,y) & \cdots & \frac{\partial F_1}{\partial y_m}(x,y) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x,y) & \cdots & \frac{\partial F_m}{\partial x_n}(x,y) & \frac{\partial F_m}{\partial y_1}(x,y) & \cdots & \frac{\partial F_m}{\partial y_m}(x,y) \end{bmatrix}$$

- $\underset{m\times n}{d_x F(x,y)} = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(x,y) & \cdots & \frac{\partial F_1}{\partial x_n}(x,y) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x,y) & \cdots & \frac{\partial F_m}{\partial x_n}(x,y) \end{bmatrix}$, $\underset{m\times m}{d_y F(x,y)} = \begin{bmatrix} \frac{\partial F_1}{\partial y_1}(x,y) & \cdots & \frac{\partial F_1}{\partial y_m}(x,y) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1}(x,y) & \cdots & \frac{\partial F_m}{\partial y_m}(x,y) \end{bmatrix}$.
- $d_y F(x_0, y_0)$ is invertible iff $L_2(v) = dF(x_0,y_0)\begin{bmatrix} 0 \\ v \end{bmatrix} = dF(x_0, y_0)v$ is bijective.
- $\begin{bmatrix} I & 0 \\ d_x F & d_y F \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -(d_y F)^{-1}(d_x F) & (d_y F)^{-1} \end{bmatrix}$
- Proof 1): by reducing the implicit function theorem to the inversion theorem.

  $H(x,y) = (x, F(x,y)) : \Omega \to \mathbb{R}^n \times \mathbb{R}^m$ is also C^1. $dH(x,y) = \begin{bmatrix} I_{n\times n} & 0_{n\times m} \\ \underset{m\times n}{d_x F(x,y)} & \underset{m\times m}{d_y F(x,y)} \end{bmatrix}$.

  $(dH(x_0,y_0))^{-1} = \begin{bmatrix} I & 0 \\ -(d_y F(x_0,y_0))^{-1} d_x F(x_0,y_0) & (d_y F(x_0,y_0))^{-1} \end{bmatrix}$, exists because $(d_y F(x_0,y_0))^{-1}$ exists.

  By inversion mapping theorem, have U, open neighborhood of (x_0, y_0), $H|_U : U \xrightarrow[\text{onto}]{1:1} V$, and $G = (H|_U)^{-1} : \underset{\mathbb{R}^n\times\mathbb{R}^m}{V} \to \underset{\mathbb{R}^n\times\mathbb{R}^m}{U}$ C^1. Define $j_1 : V \to \mathbb{R}^n$, $j_2 : V \to \mathbb{R}^m$, both C^1 by $G(x,y) = \begin{pmatrix} j_1(x,y) \\ j_2(x,y) \end{pmatrix}$. Note that

  (I) V is an open neighborhood of $H(x_0,y_0) = (x_0, 0)$.

(II) $\forall(x,y)\in V\ \ H\circ G(x,y)=(x,y)$. But $H\circ G(x,y) = H\begin{pmatrix} j_1(x,y) \\ j_2(x,y)\end{pmatrix} =$

$\begin{pmatrix} j_1(x,y) \\ F(j_1(x,y),j_2(x,y))\end{pmatrix}$. Hence, $j_1(x,y)=x$, and $F(j_1(x,y),j_2(x,y))=y \Rightarrow$
$F(x,j_2(x,y))=y$.

(III) $\forall(x,y)\in U\ \ G\circ H(x,y)=(x,y)$. But $G\circ H(x,y) = \begin{pmatrix} j_1(x,F(x,y)) \\ j_2(x,F(x,y))\end{pmatrix} =$

$\begin{pmatrix} x \\ j_2(x,F(x,y))\end{pmatrix}$. Hence, $j_2(x,F(x,y))=y$. Since $(x_0,y_0)\in U$, we have
$j_2(x_0,F(x_0,y_0))=j_2(x_0,0)=y_0$.

Let $W=\{x\in\mathbb{R}^n,(x,0)\in V\}$. Because V is an open neighborhood of $(x_0,0)$, W is an open neighborhood of x_0.

Let $f(x)=j_2(x,0):W\to\mathbb{R}^m$ C^1. $\forall x\in W$, we have $(x,0)\in V$; hence
$F(x,j_2(x,0))=0$. So, $F(x,f(x))=F(x,j_2(x,0))=0$. Also, from (III), we have
$f(x_0)=j_2(x_0,0)=y_0$.

- Proof 2) If $(x,y)\in U$, then by (III), $j_2(x,F(x,y))=y$. If, in addition $F(x,y)=0$, then $j_2(x,0)=y$. Hence, $f(x)=y$.
- Proof 3) Define $K(x)=(x,f(x)):W\to\mathbb{R}^n\times\mathbb{R}^m$. C^1. By 1) $\forall x\in W$
$F\circ K(x)=F(x,f(x))=0$. Hence, $d(F\circ K)(x)=dF(K(x))dK(x)=0$.
$\begin{bmatrix} d_xF(K(x)) & d_yF(K(x))\end{bmatrix}\begin{bmatrix} I \\ df(x)\end{bmatrix}=d_xF(K(x))+d_yF(K(x))df(x)=0$.
  - $d_yF(K(x_0))=d_yF(x_0,y_0)$ is invertible. $\exists e>0$ $|d_yF(K(x))-F(K(x_0))|<e \Rightarrow$ $d_yF(K(x))$ invertible. $F\circ K$ is $C^1\Rightarrow d_yF(K(x))$ is continuous. $\exists r>0$ $|x-x_0|<r\Rightarrow|d_yF(K(x))-F(K(x_0))|<e$.
- Regarding $df(x)=-(d_yF(x,f(x)))^{-1}d_xF(x,f(x))$,
  - if $y\in\mathbb{R}$, we have $\dfrac{\partial f}{\partial x_i}(x,y)=\dfrac{\dfrac{\partial F}{\partial x_i}(x,y)}{\dfrac{\partial F}{\partial y}(x,y)}$.

- Example: For motivation, with abuse of notation: Let $F(x,y) = y^2x_1 + 5x_2^2y + x_1x_2y^3$ $= F(1,1) = 7$. Then, $\frac{\partial F}{\partial x_1}(x, y(x)) = y^2 + x_1(2y)\frac{\partial y}{\partial x_1} + 5x_2^2\frac{\partial y}{\partial x_1} + x_2y^3 + x_1x_2 3y^2\frac{\partial y}{\partial x_1} = 0$. Hence, $\frac{\partial y}{\partial x_1}(x) = -\frac{y^2 + x_2y^3}{x_1(2y) + 5x_2^2 + x_1x_2 3y^2}$. Note that $\frac{\partial F}{\partial x_1}(x,y) = y^2 + x_2y^3 = 0$ and $\frac{\partial F}{\partial y}(x,y) = 2yx_1 + 5x_2^2 + x_1x_2 3y^2$.

- If $V \subset \mathbb{R}^n \times \mathbb{R}^m$ is an open neighborhood of (x_0, y_0), then $W = \{x \in \mathbb{R}^n, (x, y_0) \in V\}$ is an open neighborhood of x_0.

  Proof. First, because $(x_0, y_0) \in V$, we have $x_0 \in W$. Let $x' \in W$. Then $(x', y_0) \in V$. V is open; thus, $\exists d > 0$ $|(x,y) - (x', y_0)| < d \Rightarrow (x,y) \in V$. Hence, $|x - x'| < d \Rightarrow$ $|(x, y_0) - (x', y_0)| = |x - x'| < d \Rightarrow (x, y_0) \in V \Rightarrow x \in W$.

- **Implicit Function Theorem**(2: Strichartz): Let $F(x,y)$ be a C^1 function defined in a neighborhood of $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$, taking values in $\mathbb{R}^m$, with $F(x_0, y_0) = c$. Then if $d_y F(x_0, y_0)$ is invertible,

  1) there exists a neighborhood W of x_0 and a C^1 function $f: W \to \mathbb{R}^m$ such that $f(x_0) = y_0$ and $F(x, f(x)) = c$ $\forall x \in W$.
  2) Furthermore, f is unique in that there exists a neighborhood $f(W)$ of y_0 such that there is only one solution $y' \in f(W)$ of $F(x, y') = c$, namely $y' = f(x)$.
  3) Finally, the differential of f can be computed by implicit differentiation as $df(x) = -\left(d_y F(x, f(x))\right)^{-1} d_x F(x, f(x))$.

  Proof. Let $G(x,y) = F(x,y) - F(x_0, y_0)$. Then, $G(x_0, y_0) = 0$, and $d_y G(x_0, y_0) = d_y F(x_0, y_0)$. Thus, if $d_y F(x_0, y_0)$ is invertible, $d_y G(x_0, y_0)$ is also invertible, and by the implicit function theorem, we have neighborhood W of x_0 and a unique C^1 function f such that $f(x_0) = y_0$, and $\forall x \in W$ $F(x, f(x)) = F(x_0, y_0)$.

- **Inverse Function Theorem** (Strichartz): Let f be a C^1 function defined in a neighborhood of c in $\mathbb{R}^n$ taking values in $\mathbb{R}^n$. If $\underset{n \times n}{df(c)}$ is invertible, then there exists a neighborhood V of $f(c)$ and a C^1 function $g: V \to \mathbb{R}^n$ such that

  1) $f(g(y)) = y$ $\forall y \in V$.
  2) Furthermore, g maps V one-to-one onto a neighborhood U of c and $g(f(x)) = x$ $\forall x \in U$.

3) The function g is unique in that for any y in V, there is only one x' in U with $f(x') = y$ namely $x' = g(y)$.

4) Finally, $dg(y) = [df(x)]^{-1}$ if $f(x) = y$.

- Inverse function theorem is a special case of the implicit theorem.
  Given a C^1 function g and $\underset{n\times n}{dg(y_0)}$ is invertible. Then, let $F(x, y) = g(y) - x$, and $x_0 = g(y_0)$. Note that $F(x_0, y_0) = g(y_0) - x_0 = 0$. $d_y F(x_0, y_0) = dg(y_0)$ invertible. Then $\exists W$ open neighborhood of $x_0 = g(y_0)$, and a unique function $f = g^{-1}$ of class C^1, $f : W \to \mathbb{R}^m$, $f(x_0) = y_0$, and $\forall x \in W$ $g(f(x)) = x$. Also, $\exists U$ open neighborhood of (x_0, y_0) such that if $(x, y) \in U$ and $g(y) = x$ then $x \in W$ and $y = f(x)$. Moreover, $\exists r > 0$ $|x - x_0| < r \Rightarrow df(x) = -(dg(f(x)))^{-1}(-I) = (dg(f(x)))^{-1}$.
- Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a C^1 function. Then, the restriction of f to any open set of $\mathbb{R}^2$ is not injective.

Summary

- Let open $\Omega \subset \mathbb{R}^n$, $f:\Omega \to \mathbb{R}^m$ is C^1, $c \in \Omega$.
  - **The injective mapping theorem**:
    $df(c)$ is injective $\Rightarrow$ $\exists d>0$ such that $f|_{B_d(c)}$ is injective.
    1) $f|_{B_d(c)}$ is injective
    2) $\forall x \in B_d(c)$ $df(x)$ is injective.
    3) $g=\left(f|_{B_d(c)}\right)^{-1}: f\left(B_d(c)\right) \xrightarrow[\text{onto}]{1:1} B_d(c)$ is unique, and uniformly continuous.
    4) $\exists r>0$ $\forall x_1 \forall x_2$ $x_1,x_2 \in B_d(c) \Rightarrow \left|f(x_2)-f(x_1)\right| \geq \frac{r}{2}\left|x_2-x_1\right|$.
  - The **surjective mapping theorem**:
    $df(c)$ is surjective $\Rightarrow$ $\exists e_1, e_2>0$ such that $B_{e_1}(c) \subset \Omega$ and $B_{e_2}\left(f(c)\right) \subset f\left(B_{e_1}(c)\right)$.
  - **Open mapping theorem**.
    $\forall x \in \Omega$ $df(x)$ is surjective. $\Rightarrow$ f is an open mapping.
- **Inversion (Mapping) Theorem**: Let open $\Omega \subset \mathbb{R}^n$, $f:\Omega \to \mathbb{R}^n$ is C^1, $c \in \Omega$. $\underset{n\times n}{df(c)}$ is bijective, then $\exists U$ open neighborhood of c such that
  1) $V = f(U)$ is an open neighborhood of $f(c)$.
  2) $f|_U : U \to V$ is bijective.
  3) $g=\left(f|_U\right)^{-1}: V \to U$ is C^1.
  4) $\forall y \in V$ $dg(y)=\left[df\left(g(y)\right)\right]^{-1}$.
- **Implicit Function Theorem**: Let $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ open. $F:\Omega \to \mathbb{R}^m$ C^1. $(x_0,y_0) \in \Omega$, $F(x_0,y_0)=0$, $d_y F(x_0,y_0)$ is invertible, then
  1) $\exists W$ open neighborhood of x_0, and a unique function f of class C^1, $f:W \to \mathbb{R}^m$, $f(x_0)=y_0$, and $\forall x \in W$ $F\left(x, f(x)\right)=0$.
  2) Also, $\exists U$ open neighborhood of (x_0,y_0) such that if $(x,y) \in U$ and $F(x,y)=0$ then $x \in W$ and $y=f(x)$.
  3) Moreover, $\exists r>0$ $\left|x-x_0\right|<r \Rightarrow df(x)=-\left(d_y F\left(x, f(x)\right)\right)^{-1} d_x F\left(x, f(x)\right)$.