For any $\bar{x} = (x_1, ...x_n), \bar{y} = (y_1, ...y_n)$ we denote by $< \bar{x}, \bar{y} >$ the scalar product $< \bar{x}, \bar{y} > := \sum_{i=1}^n x_i y_i$ and define $\|\bar{x}\|^2 :=< \bar{x}, \bar{x} >$.

The Inverse Function Theorem (IFT).

Let $U \subset \mathbb{R}^n$ be an open set $u_0 \in U$ and $f : U\ \mathbb{R}^n$ a continuously differentiable map such that the linear map $D_f(u_0) : \mathbb{R}^n \to \mathbb{R}^n$ is invertible. Let $v_0 := f(u_0) \in \mathbb{R}^n$. Then there exists an open set $V \subset \mathbb{R}^n$ containing v_0 and a continuously differentiable map $g : V \to U$ such that

1)$g(v_0) = (u_0)$

2)$f \circ g(v) = v$ for all $v \in V$

3) The image $g(V) \subset \mathbb{R}^n$ of g is open.

Remark. If $U = \mathbb{R}^n$ and the map f is linear map then we can take $V = \mathbb{R}^n$ and we have $g = f^{-1}$.

Proof of the theorem is based on a the following auxiliary result.

Lemma 1. Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator such that $\|A\| < 1, \bar{x} = (x_1, ...x_n) \in \mathbb{R}^n$ and $\bar{y} = (y_1, ...y_n) := A\bar{x}$. Then for all $i, 1 \leq i \leq n$ we have $|y_i| \leq c\sqrt{n}$ where $c := max|x_j|, 1 \leq j \leq n$.

I'll leave it as an excercise.

We will also use the following result [the multivariable MVT]. Let $U \subset \mathbb{R}^n, u', u'' \in U$ be two points such that the interval $I_{u',u''} \subset \mathbb{R}^n$ between them is contained in $U, f : U \to \mathbb{R}$ a continuously differentiable function. Then there exists a point $c \in I_{u',u''}$ such that $f(u'') - f(u') = D_f(c)(u'' - u')$. This result is a special case of the taylor Remainder Theorem for $m = 0$. See the page 192 of your book or the notes.

Proof of the IFT. We first consider the special case when $D_f(u_0) = Id$. Since the set U is open, the function f is continuously differentiable and $D_f(u_0) = Id$ there exists $\epsilon > 0$ such that

a)$B_{u_0}(\epsilon) \subset U$

b)$\|D_f(u) - D_f(u_0)\| < 1/2\sqrt{n}$ and

c)$\|f(u) - f(u_0) - (u - u_0)\| \leq 1/10\|u - u_0\|$

for all $u \in B_{u_0}(\epsilon)$.

In particular we have $\|f(u) - v_0\| \geq 9/10\|u - u_0\|$ [since $v_0 = f(u_0)$]. Also we see that for any $u \in B_{u_0}(\epsilon)$ we have $\|D_f(u) - Id| < 1/2\sqrt{n}$ [since $D_f(u_0) = Id$] and therefore the linear map $D_f(u)$ is invertible.

We can write $f(u) = (f_1(u), ..., f_n(u))$ where $f_i : U \to \mathbb{R}$ are continuously differentiable functions. Moreover $D_f(u)(h) = (D_{f_1}(u)(h), ..., D_{f_n}(u)(h))$ for $h \in \mathbb{R}^n$.

Lemma 1'.Since $\|D_f(u) - Id| < 1/2$ we have

$$\partial f/\partial x_i f_i(u) > 1 - 1/2 = 1/2$$

[Please show how this inequality follows from Lemma 1]

Let $B = B_{u_0}(\epsilon), \bar{B}$ be the closure of the open ball B. That is $\bar{B} = \{u \in U\}$ such that $\||u - u_0\| \leq \epsilon$. Then $\bar{B}$ is a closed and bounded subset of $\mathbb{R}^n$. Therefore $\bar{B}$ is compact.

Remark. $\bar{B} - B = \{u \in \bar{B}|\||u - u_0\| = \epsilon$.

Let $V := B_{v_0}(\epsilon/10)$. We will construct now a function $g : V \to B$. To define a function $g : V \to U_1$ we have to define for any $v \in V$ an element $g(v) \in B$. To define $g(v) \in B$ consider a function F_v on $\bar{B}$ by $F_v(u) := \|v - f(u)\|^2, u \in B$.

Lemma 2. The function F_v on $\bar{B}$ is continuous.

Proof ?

Since the function F_v on $\bar{B}$ is continuous and $\bar{B}$ is compact is compact there exists $u_v \in \bar{B}$ such that $F_v(u_v) \leq F_v(u)$ for all $u \in B$.

Proposition 1. $u_v \in B$.

Proof of Proposition 1. Assume that u_v does not belong to B. Then $\|u_v - u_0\| = \epsilon$. By the definition we have $F_v(u_v) = \|v - f(u_v)\|^2$. But $\|v - f(u_v)\| \geq \|f(u_v) - v_0\| - \|v - v_0\| \geq 9/10\|u_v - u_0\| - \epsilon/10$ since $v \in V := B_{v_0}(\epsilon/10)$. By our assumption u_v does not belong to B. So $\|u_v - u_0\| = \epsilon$ and we see that $\|v - f(u_v)\| \geq 4/5\epsilon$ and therefore $F_v(u_v) \geq (4/5\epsilon)^2$. On the other hand, we have $F_v(u_0) = \|v - f(u_0)\|^2 = \|v - v_0\|^2 = (\epsilon/10)^2$. So $F_v(u_0) < F_v(u_v)$. This contrudiction proves the Proposition 1.

Proposition 2. $F_v(u_v) = 0$

Proof of Proposition 2. Since the function f is continuously differentiable we see [Chain Rule] that the function F_v on B is continuously differentiable and the linear map $D_{F_v}(u) : \mathbb{R}^n \to \mathbb{R}$ is given by

$$D_{F_v}(u)(h) =< 2D_f(u)(h), (f(u) - v) >$$

[Please check].

Since the function F_v reaches it's minimum at $u_v \in B$ we have $D_{F_v}(u_v) = 0$ [Please explain]

Since the the linear map $D_f(u) : \mathbb{R}^n \to \mathbb{R}^n$ is invertible we see that $< w, (f(u_v) - v) >= 0$ for all $w \in \mathbb{R}^n$. Therefore $f(u_v) - v = 0$. So $v = f(u_v)$ and $F_v(u_v) = 0$. Proposition 2 is proven

Proposition 3. For any $v \in V$ there exists unique $g(v) \in B$ such that $f(g(v)) = v$.

We have seen that there exists a point $u' \in B$ such that $v = f(u')$. So we have only to prove the uniqueness of such a point. Let $u'' \in B$ be a point such that $v = f(u'')$. We assume that $u' \neq u''$ and come to a contrucdiction.

So assume that $u' \neq u'', u' = (x_1', ...x_n'), u'' = (x_1'', ...x_n'')$. If $u' \neq u''$ there exists $i, 1 \leq i \leq n$ such that $u_i' \neq u_i''$. We can assume by choosing an appropriate i that $|u_i' - u_i''| \geq |u_j' - u_j''|$ for all $j, 1 \leq j \leq n$.

As follows from the multivariable MVT there exists $c \in I_{u',u''} \subset B$ such that $f_i(u'') - f_i(u') = D_{f_i}(c)(u'' - u')$. Since $D_f(c) = Id + (D_f(c) - Id)$ and by the choice of ϵ we have $\|(D_f(c) - Id)\| < 1/2\sqrt{n}$ it follows from Lemma 1' that $|f_i(u'') - f_i(u')| > 1/2|u_i'' - u_i'|$ [Please explain how to derive this inequality from Lemma 1']. Proposition 3 is proven.

It follow from Proposition 3 we can define a function $g : V \to B$ where for any $v \in V, g(v) \in B$ is the unique point such that $f(g(v)) = v$.

Proposition 4. The function $g : V \to B$ is continuously differentiable at any point $y \in V$ and $D_g(y) = D_f(x)^{-1}$ where $x := g(y)$.

Proof. The proof is very similar the the proof in the one-dimensional case which we have done in class. It is also well written in your book on the page 290-291.

Now we drop the assumption that $D_f(u_0) = Id$. Let $f : U \to \mathbb{R}^n$ be a continuously differentiable map such that the linear map $D_f(u_0) : \mathbb{R}^n \to \mathbb{R}^n$ is invertible. Define $A := (D_f(u_0))^{-1}$. Then A is a linear map from $\mathbb{R}^n$ to itself. Define $\tilde{f} := A \circ f$. Then $\tilde{f} : U \to \mathbb{R}^n$ is a continuously differentiable map and as follows from the Chain rule $D_{\tilde{f}}(u_0) = Id$. Therefore it follows from Propositions 3 and 4 that there exists an open set $\tilde{V}$ containing $\tilde{f}(u_0)$ and a continuously differentiable function $\tilde{g} : \tilde{V} \to B$ such that $\tilde{f}(\tilde{g}(\tilde{v})) = \tilde{v}$ from all $\tilde{v} \in \tilde{V}$. Consider now an open (?) set $V := D_f(u_0)(\tilde{V}) \subset \mathbb{R}^n$. Then $f(u_0) \in V$ and the restriction of the linear map A to V defines a map $V \to \tilde{V}$ which we also denote by A. Then the function $g := \tilde{g} \circ A : V \to B$ is continuously differentiable and $f(g(v)) = v$ for all $v \in V$.

To finish the proof of the IFT we have to show that the image $g(V)$ is an open subset of $\mathbb{R}^n$. But this follows from the Proposition 3 if you apply it to the map g