Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

$\mathbb{R}^n = \{(x_1, \dots, x_n) |\ x_i \in \mathbb{R},\ i = 1, \dots, n\}$

$\mathbb{R}^n$ is a vector space with standard basis $\{\vec{e}_1, \dots, \vec{e}_n\}$ where $\vec{e}_i = < 0, 0, 0, 1, 0, \dots, 0 >$ (1 in the i^{th} place). The standard norm on $\mathbb{R}^n$ is given by:

$$\|\vec{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}, \qquad \text{where } \vec{x} = < x_1, \dots, x_n >.$$

We can define a distance on $\mathbb{R}^n$ by:

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}\ .$$

Def. If $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, we say that f is **continuous at $\vec{a} \in A$** if for all $\epsilon > 0$ there exists a $\delta > 0$ such that if $\|\vec{x} - \vec{a}\| < \delta$, then $\|f(\vec{x}) - f(\vec{a})\| < \epsilon$.

Def. If $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$ and $\vec{a} \in A$, then we define the ***i^{th}* partial derivative** of f at $\vec{a}$ as:

$$\frac{\partial f}{\partial x_i}(\vec{a}) = \lim_{h \to 0} \frac{f(a_1, \dots, a_i + h, \dots, a_n) - f(a_1, \dots, a_n)}{h}$$

as long as the limit exists.

Def. If $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, we say that f is **differentiable at $\vec{a} \in A$** if there exists a linear transformation $\lambda: \mathbb{R}^n \to \mathbb{R}^m$ such that:

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

In this case, we say $Df(\vec{a}) = \lambda$.

Let $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, then we can write:

$$f(x_1, \dots, x_n) = \big(f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_m(x_1, x_2, \dots, x_n)\big)$$

where $f_i: \mathbb{R}^n \to \mathbb{R}$.

Theorem: If $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec{a} \in \mathbb{R}^n$, then $\frac{\partial f_i}{\partial x_j}(\vec{a})$ exists for $1 \le i \le m$, $1 \le j \le n$, and

$$Df(\vec{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

where $\frac{\partial f_i}{\partial x_j}$ is evaluated at $\vec{a}$.

Def. $Df(\vec{a})$ is called the **Jacobian matrix** of f at $\vec{a}$. So if $Df(\vec{a})$ exists, then all of the partial derivatives, $\frac{\partial f_i}{\partial x_j}$, exist at $\vec{a}$. The converse is not true: all of $\frac{\partial f_i}{\partial x_j}$ existing at $\vec{a}$ does not imply $Df(\vec{a})$ exists.

Theorem (Chain Rule): If $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\vec{a} \in \mathbb{R}^n$, and $g: \mathbb{R}^m \to \mathbb{R}^p$ is differentiable at $f(\vec{a})$, then $g \circ f: \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at $\vec{a}$ and $D(g \circ f)(\vec{a}) = Dg\big(f(\vec{a})\big) \circ Df(\vec{a})$.

Def. A **change of coordinates** on an open set, $U \subseteq \mathbb{R}^n$, is a differentiable map $g: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ such that $det\big(Dg(\vec{x})\big) \neq 0$ for $\vec{x} \in U$.

Ex. Suppose $g: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a change of coordinates on U. Thus:

$$g(x_1, .., x_n) = (g_1(x_1, .., x_n), \dots, g_n(x_1, .., x_n))$$

$$\begin{aligned} \bar{x}_1 &= g_1(x_1, .., x_n) \\ &\vdots \\ \bar{x}_n &= g_n(x_1, .., x_n). \end{aligned}$$

Given $f: \mathbb{R}^n \to \mathbb{R}$, find a relationship between $\frac{\partial f}{\partial x_i}$ and $\frac{\partial f}{\partial \bar{x}_1}, \dots, \frac{\partial f}{\partial \bar{x}_n}$.

By the Chain Rule, since $f \circ g: \mathbb{R}^n \to \mathbb{R}$:

$$D(f \circ g)(x_1, \dots, x_n) = Df\big(g(x_1, \dots, x_n)\big) \circ Dg(x_1, \dots, x_n)$$

$$\begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial \bar{x}_1} & \cdots & \frac{\partial f}{\partial \bar{x}_n} \end{pmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{pmatrix}$$

$$\begin{pmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial \bar{x}_1} & \cdots & \frac{\partial f}{\partial \bar{x}_n} \end{pmatrix} \begin{pmatrix} \frac{\partial \bar{x}_1}{\partial x_1} & \cdots & \frac{\partial \bar{x}_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \bar{x}_n}{\partial x_1} & \cdots & \frac{\partial \bar{x}_n}{\partial x_n} \end{pmatrix}$$

$$\boxed{\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial \bar{x}_1}\frac{\partial \bar{x}_1}{\partial x_i} + \cdots + \frac{\partial f}{\partial \bar{x}_n}\frac{\partial \bar{x}_n}{\partial x_i}}$$

Ex. Let $\bar{x}_1 = x_1 \cos x_2$; $\bar{x}_2 = x_1 \sin x_2$; $f(\bar{x}_1, \bar{x}_2) = (\bar{x}_1^2 + \bar{x}_2^2) + \bar{x}_1\bar{x}_2$.
Calculate $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}$ by the Chain Rule and directly and show they are equal.

Chain Rule: $\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial \bar{x}_1}\frac{\partial \bar{x}_1}{\partial x_1} + \frac{\partial f}{\partial \bar{x}_2}\frac{\partial \bar{x}_2}{\partial x_1}$; $\frac{\partial f}{\partial x_2} = \frac{\partial f}{\partial \bar{x}_1}\frac{\partial \bar{x}_1}{\partial x_2} + \frac{\partial f}{\partial \bar{x}_2}\frac{\partial \bar{x}_2}{\partial x_2}$

$$\frac{\partial \bar{x}_1}{\partial x_1} = \cos x_2; \qquad \frac{\partial \bar{x}_2}{\partial x_1} = \sin x_2$$

$$\frac{\partial \bar{x}_1}{\partial x_2} = -x_1 \sin x_2; \qquad \frac{\partial \bar{x}_2}{\partial x_2} = x_1 \cos x_2$$

$$\frac{\partial f}{\partial \bar{x}_1} = 2\bar{x}_1 + \bar{x}_2; \qquad \frac{\partial f}{\partial \bar{x}_2} = 2\bar{x}_2 + \bar{x}_1.$$

$$\frac{\partial f}{\partial x_1} = (2\bar{x}_1 + \bar{x}_2)(\cos x_2) + (2\bar{x}_2 + \bar{x}_1)(\sin x_2)$$

$$= (2x_1 \cos x_2 + x_1 \sin x_2)(\cos x_2) + (2x_1 \sin x_2 + x_1 \cos x_2)(\sin x_2)$$

$$= 2x_1(\cos^2 x_2 + \sin^2 x_2) + 2x_1(\sin x_2)(\cos x_2)$$

$$= 2x_1 + 2x_1(\sin x_2)(\cos x_2).$$

$$\frac{\partial f}{\partial x_2} = (2\bar{x}_1 + \bar{x}_2)(-x_1 \sin x_2) + (2\bar{x}_2 + \bar{x}_1)(x_1 \cos x_2)$$

$$= (2x_1 \cos x_2 + x_1 \sin x_2)(-x_1 \sin x_2)$$
$$+(2x_1 \sin x_2 + x_1 \cos x_2)(x_1 \cos x_2)$$

$$= -x_1^2 \sin^2 x_2 + x_1^2 \cos^2 x_2 \,.$$

Directly: $f(\bar{x}_1, \bar{x}_2) = \bar{x}_1^2 + \bar{x}_2^2 + \bar{x}_1\bar{x}_2 = x_1^2 + x_1^2 \sin x_2 \cos x_2$

$$\frac{\partial f}{\partial x_1} = 2x_1 + 2x_1 \sin x_2 \cos x_2$$

$$\frac{\partial f}{\partial x_2} = x_1^2(-\sin^2 x_2 + \cos^2 x_2).$$

Ex. Let $u = x - 3y$ and $v = x + 3y$. Let $T: \mathbb{R}^2 \to \mathbb{R}$ be a smooth function. Find $\frac{\partial^2 T}{\partial u \partial v}$ only in terms of derivatives of T with respect to x and y.

By the chain rule:

$$\frac{\partial T}{\partial v} = \frac{\partial T}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial T}{\partial y}\frac{\partial y}{\partial v}.$$

So we need x and y in terms of u and v so we can calculate $\frac{\partial x}{\partial v}$ and $\frac{\partial y}{\partial v}$.

$$\begin{aligned} u &= x - 3y \\ \underline{v} &\underline{= x + 3y} \\ u + v &= 2x \end{aligned} \qquad \begin{aligned} u &= x - 3y \\ \underline{v} &\underline{= x + 3y} \\ u - v &= -6y \end{aligned}$$

Or $x = \frac{1}{2}(u + v)$ or $y = \frac{1}{6}(v - u)$.

$$\frac{\partial x}{\partial u} = \frac{1}{2} \qquad \frac{\partial x}{\partial v} = \frac{1}{2} \qquad \frac{\partial y}{\partial u} = -\frac{1}{6} \qquad \frac{\partial y}{\partial v} = \frac{1}{6}.$$

$$\frac{\partial T}{\partial v} = \frac{\partial T}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial T}{\partial y}\frac{\partial y}{\partial v} = \frac{1}{2}\frac{\partial T}{\partial x} + \frac{1}{6}\frac{\partial T}{\partial y}$$

$$\frac{\partial^2 T}{\partial u \partial v} = \frac{\partial}{\partial u}\left[\frac{\partial T}{\partial v}\right] = \frac{\partial}{\partial u}\left[\frac{1}{2}\frac{\partial T}{\partial x} + \frac{1}{6}\frac{\partial T}{\partial y}\right] = \frac{1}{2}\left[\frac{\partial}{\partial u}\frac{\partial T}{\partial x}\right] + \frac{1}{6}\left[\frac{\partial}{\partial u}\frac{\partial T}{\partial y}\right]$$

$$= \frac{1}{2}\left[\frac{\partial^2 T}{\partial x^2}\frac{\partial x}{\partial u} + \frac{\partial^2 T}{\partial y \partial x}\frac{\partial y}{\partial u}\right] + \frac{1}{6}\left[\frac{\partial^2 T}{\partial x \partial y}\frac{\partial x}{\partial u} + \frac{\partial^2 T}{\partial y^2}\frac{\partial y}{\partial u}\right]$$

$$= \frac{1}{2}\left[\frac{\partial^2 T}{\partial x^2}\left(\frac{1}{2}\right) + \frac{\partial^2 T}{\partial y \partial x}\left(-\frac{1}{6}\right)\right] + \frac{1}{6}\left[\frac{\partial^2 T}{\partial x \partial y}\left(\frac{1}{2}\right) + \frac{\partial^2 T}{\partial y^2}\left(-\frac{1}{6}\right)\right]$$

$$\frac{\partial^2 T}{\partial u \partial v} = \frac{1}{4}\frac{\partial^2 T}{\partial x^2} - \frac{1}{36}\frac{\partial^2 T}{\partial y^2}.$$

Directional Derivatives:

Def. Let $F: U \subseteq \mathbb{R}^n \to \mathbb{R}$, $\vec{a} \in U$, and $\vec{u}$ be a unit vector in $\mathbb{R}^n$. The **directional derivative** of F at $\vec{a}$ in the direction $\vec{u}$ is:

$$D_{\vec{u}}F(\vec{a}) = \lim_{h \to 0} \frac{F(\vec{a} + h\vec{u}) - F(\vec{a})}{h}$$

when the limit exists.

Notice:

$$D_{\vec{u}}F(\vec{a}) = \frac{d}{dt}\left(F(\vec{a} + t\vec{u})\right)\Big|_{t=0}$$

since if $g(t) = F(\vec{a} + t\vec{u})$, then:

$$g'(0) = \lim_{h \to 0} \frac{g(0+h) - g(0)}{h} = \lim_{h \to 0} \frac{F(\vec{a}+h\vec{u}) - F(\vec{a})}{h}.$$

Also notice by the Chain Rule:

$$\frac{d}{dt}\big(F(\vec{a}+t\vec{u})\big) = \frac{\partial F}{\partial x_1}\frac{\partial x_1}{\partial t} + \cdots + \frac{\partial F}{\partial x_n}\frac{\partial x_n}{\partial t}\,.$$

So if $x_1 = a_1 + tu_1 \;\Rightarrow\; \dfrac{dx_1}{dt} = u_1$

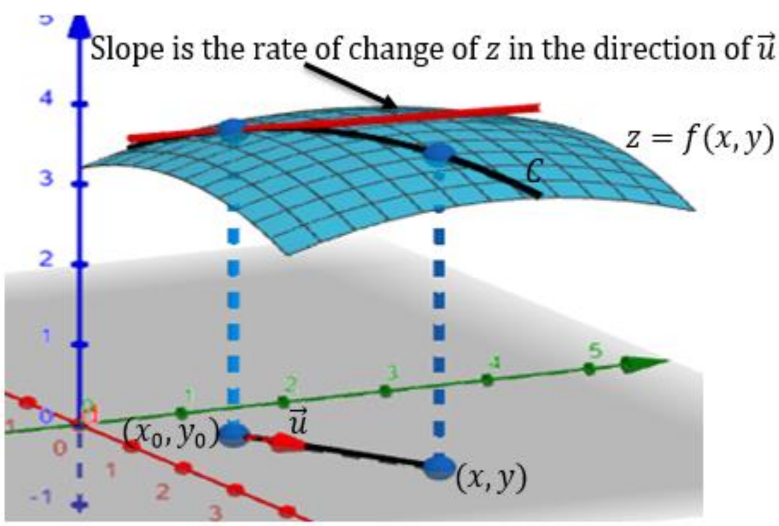
$\vdots$

$x_n = a_n + tu_n \;\Rightarrow\; \dfrac{dx_n}{dt} = u_n$ then:

$$D_{\vec{u}}F(\vec{a}) = \frac{d}{dt}\big(F(\vec{a}+t\vec{u})\big)\Big|_{t=0} = \nabla F\cdot\vec{u}\,.$$

So $D_{\vec{u}}F(\vec{a})$ **is the rate of change of the value of F in the direction of $\vec{u}$ at $\vec{a}$.**

In the case of a function $z = f(x,y)$ we have:



Ex. Find the directional derivative of $F(x_1, x_2, x_3) = x_1^2 + x_2^2 x_3$ at the point $(3, 2, 1)$ in the direction of $\vec{u} = \frac{1}{\sqrt{21}}(-1, -4, 2)$.

$$D_{\vec{u}}F(3,2,1) = \nabla F(3,2,1) \cdot \vec{u}$$

$$\nabla F = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \frac{\partial F}{\partial x_3}\right) = (2x_1, 2x_2x_3, x_2^2)$$

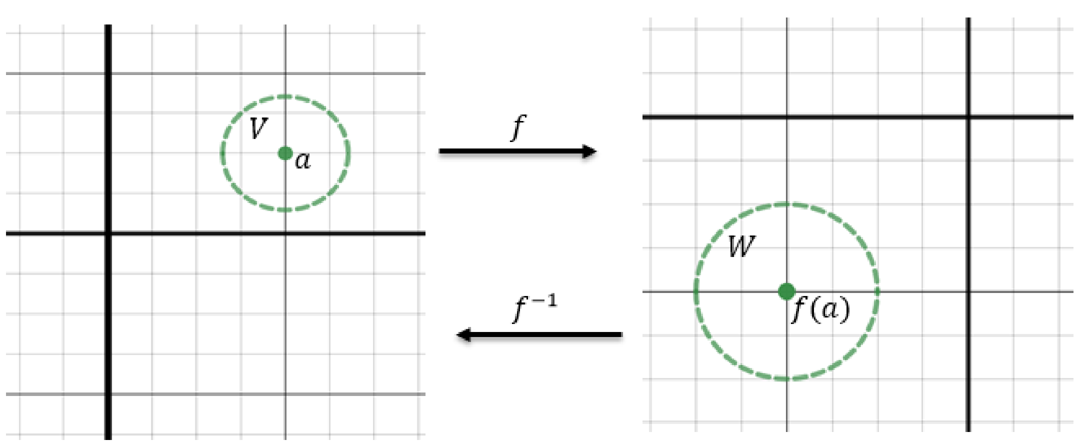
$$\nabla F(3,2,1) = (6,4,4)$$

$$D_{\vec{u}}F(3,2,1) = (6,4,4) \cdot \frac{1}{\sqrt{21}}(-1,-4,2) = -\frac{14}{\sqrt{21}}.$$

Note: If $DF(\vec{a})$ exists, then so do all of the partial derivatives of F and hence ∇F. Thus, the directional derivative of F in the direction of $\vec{u}$ at $\vec{a}$ also exists, since $D_{\vec{u}}F(\vec{a}) = \nabla F \cdot \vec{u}$.

Inverse Function Theorem: Suppose that $f: \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable in an open set U, containing $\vec{a}$ (i.e. $Df(\vec{x})$ exists for all $\vec{x} \in U$ and $\frac{\partial f_i}{\partial x_j}$ is continuous at $\vec{x} \in U$ for $i = 1, \dots, n$) and $det\big(Df(\vec{a})\big) \neq 0$, then there exists an open set , V, containing $\vec{a}$ and an open set, W, containing $f(\vec{a})$ such that $f: V \to W$ has a continuous inverse $f^{-1}: W \to V$, which is continuously differentiable for $y \in W$ and satisfies:

$$(f^{-1})'(y) = \left[f'\big(f^{-1}(y)\big)\right]^{-1}.$$



Note 1: If $det\big(Df(\vec{a})\big) = 0$, f might still have a continuous inverse, however its inverse is not differentiable. For example, $f(x) = x^3$ near $x = 0$.

Note 2: The Inverse Function Theorem only guarantees a local inverse. In fact, f can have a local inverse at every point and not have a global inverse. For example, $f(x, y) = (e^x \cos y\,, e^x \sin y)$.

Ex. Suppose $f(x_1, x_2) = (x_1 \cos x_2\,, x_1 \sin x_2)\,;\quad 0 < x_1\,,\ \ 0 \leq x_2 < 2\pi,$ so $\bar{x}_1 = x_1 \cos x_2,\ \ \bar{x}_2 = x_1 \sin x_2$. Find $D(f^{-1})$.

$$Df(x_1, x_2) = \begin{pmatrix} \dfrac{\partial \bar{x}_1}{\partial x_1} & \dfrac{\partial \bar{x}_1}{\partial x_2} \\ \dfrac{\partial \bar{x}_2}{\partial x_1} & \dfrac{\partial \bar{x}_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \cos x_2 & -x_1 \sin x_2 \\ \sin x_2 & x_1 \cos x_2 \end{pmatrix}$$

$$\det\big(Df(x_1, x_2)\big) = x_1 \cos^2 x_2 + x_1 \sin^2 x_2 = x_1 \neq 0.$$

So by the Inverse Function Theorem:

$$D(f^{-1})(\bar{x}_1, \bar{x}_2) = [Df(x_1, x_2)]^{-1}\,.$$

Recall for a 2 by 2 matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \Longrightarrow \quad A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Thus we have:

$$D(f^{-1})(\bar{x}_1, \bar{x}_2) = \frac{1}{x_1} \begin{pmatrix} x_1 \cos x_2 & x_1 \sin x_2 \\ -\sin x_2 & \cos x_2 \end{pmatrix}.$$

In terms of $\bar{x}_1, \bar{x}_2$:

$$\bar{x}_1 = x_1 \cos x_2 \qquad \bar{x}_2 = x_1 \sin x_2 \qquad x_1 = \sqrt{\bar{x}_1^2 + \bar{x}_2^2}\,.$$

So we can write:

$$D(f^{-1})(\bar{x}_1, \bar{x}_2) = \frac{1}{\sqrt{\bar{x}_1^2 + \bar{x}_2^2}} \begin{pmatrix} \bar{x}_1 & \bar{x}_2 \\ -\dfrac{\bar{x}_2}{x_1} & \dfrac{\bar{x}_1}{x_1} \end{pmatrix} = \begin{pmatrix} \dfrac{\bar{x}_1}{\sqrt{\bar{x}_1^2 + \bar{x}_2^2}} & \dfrac{\bar{x}_2}{\sqrt{\bar{x}_1^2 + \bar{x}_2^2}} \\ -\dfrac{\bar{x}_2}{\bar{x}_1^2 + \bar{x}_2^2} & \dfrac{\bar{x}_1}{\bar{x}_1^2 + \bar{x}_2^2} \end{pmatrix}.$$

Ex. For a general change of coordinates:

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

$$g(x_1, \dots, x_n) = \big(g_1(x_1, \dots, x_n), \dots, g_n(x_1, \dots, x_n)\big)$$

$$\begin{aligned} \bar{x}_1 &= g_1(x_1, \dots, x_n) = \bar{x}_1(x_1, \dots, x_n) \\ &\vdots \qquad\qquad\qquad\qquad \vdots \\ \bar{x}_n &= g_n(x_1, \dots, x_n) = \bar{x}_n(x_1, \dots, x_n). \end{aligned}$$

The inverse map, g^{-1}, takes $\bar{x}_1, \dots, \bar{x}_n$ into $x_1, \dots, x_n$. That is:

$$\begin{aligned} x_1 &= x_1(\bar{x}_1, \dots, \bar{x}_n) \\ &\vdots \\ x_n &= x_n(\bar{x}_1, \dots, \bar{x}_n). \end{aligned}$$

Thus, we have:

$$(Dg)(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial \bar{x}_1}{\partial x_1} & \cdots & \frac{\partial \bar{x}_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \bar{x}_n}{\partial x_1} & \cdots & \frac{\partial \bar{x}_n}{\partial x_n} \end{pmatrix}$$

$$(Dg^{-1})(\bar{x}_1, \dots, \bar{x}_n) = \begin{pmatrix} \frac{\partial x_1}{\partial \bar{x}_1} & \cdots & \frac{\partial x_1}{\partial \bar{x}_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial \bar{x}_1} & \cdots & \frac{\partial x_n}{\partial \bar{x}_n} \end{pmatrix}.$$

By the Inverse Function Theorem,

$$\begin{pmatrix} \frac{\partial \bar{x}_1}{\partial x_1} & \cdots & \frac{\partial \bar{x}_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \bar{x}_n}{\partial x_1} & \cdots & \frac{\partial \bar{x}_n}{\partial x_n} \end{pmatrix} \text{and} \begin{pmatrix} \frac{\partial x_1}{\partial \bar{x}_1} & \cdots & \frac{\partial x_1}{\partial \bar{x}_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial \bar{x}_1} & \cdots & \frac{\partial x_n}{\partial \bar{x}_n} \end{pmatrix}$$

are inverses of each other.

$$\begin{pmatrix} \frac{\partial x_1}{\partial \bar{x}_1} & \cdots & \frac{\partial x_1}{\partial \bar{x}_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial \bar{x}_1} & \cdots & \frac{\partial x_n}{\partial \bar{x}_n} \end{pmatrix} \begin{pmatrix} \frac{\partial \bar{x}_1}{\partial x_1} & \cdots & \frac{\partial \bar{x}_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \bar{x}_n}{\partial x_1} & \cdots & \frac{\partial \bar{x}_n}{\partial x_n} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 1 \end{pmatrix}.$$

In other words:

$\sum_{i=1}^{n} \frac{\partial x_k}{\partial \bar{x}_i} \frac{\partial \bar{x}_i}{\partial x_j} = \delta_j^k$; δ_j^k is known as the **Kronecker Delta**.