Note: \( \frac{\partial f}{\partial s_j} = \frac{\partial}{\partial s_j} \left( f(x + s_j, y) - f(x) \right) = \frac{\partial f}{\partial x_j} (x + s_j) - \frac{\partial f}{\partial x_j} (x) \)

\[ \frac{\partial f}{\partial s_j} (x) = \lim_{t \to 0} \lim_{s \to 0} \frac{1}{t} \left( \frac{\partial f}{\partial x_j} (x + s \varepsilon_j + t \varepsilon_i) - \frac{\partial f}{\partial x_j} (x + s \varepsilon_j) \right) \]

View \( \frac{\partial f}{\partial x_j} (x + s \varepsilon_j + t \varepsilon_i) \) as a function of \( t \) (and \( s \)) and define:

\[ h(t, s) = \frac{\partial f}{\partial x_j} (x + s \varepsilon_j + t \varepsilon_i) \]

Then:

\[ \frac{\partial f}{\partial x_j} (x) = \lim_{t \to 0} \lim_{s \to 0} \left( \frac{1}{t} [h(t, s) - h(t, 0)] \right) \]

Apply Mean Value Theorem:

\[ \frac{\partial f}{\partial x_j} (x) = \lim_{t \to 0} \lim_{s \to 0} \frac{1}{t} \frac{\partial h}{\partial t} (\varepsilon, \bar{s}) = \lim_{t \to 0} \lim_{s \to 0} \frac{\partial h}{\partial t} (\varepsilon, \bar{s}) \text{ where } \varepsilon \text{ lies between } 0 \text{ and } t. \]

Note:

\[ \frac{\partial h}{\partial t} (\varepsilon, \bar{s}) = \frac{\partial^2 f}{\partial x_j \partial x_i} (x + s \varepsilon_j + t \varepsilon_i) \]

As \( s \to 0, t \to 0, \varepsilon \) & \( \bar{s} \) both are forced to 0 and the result follows.

**Remark:** Important application \((n=2)\).

Let \( x = (x_1, y) \) and \( x_0 = (x_0, y_0). \) When we study Taylor series, we will see:

\[ f(x) = f(x_0) + \frac{\partial f}{\partial x}(x_0, x-x_0) + \frac{\partial f}{\partial y}(x_0, y-y_0) + \frac{1}{2} \left( \frac{\partial^2 f}{\partial x^2}(x_0, x-x_0)^2 + \right. \]

\[ \left. \frac{\partial^2 f}{\partial x \partial y}(x_0, x-x_0)(y-y_0) + \frac{\partial^2 f}{\partial y^2}(x_0, y-y_0)^2 \right) + \ldots \]

\[ f(x) - f(x_0) = \left[ f_x(x_0) f_y(x_0) \right] \left[ x-x_0 \right] + \frac{1}{2} \left[ x-x_0 \ y-y_0 \right] \left[ f_{xx}(x_0) f_{xy}(x_0) \right] \left[ x-x_0 \right] + \ldots \]

If \( x_0 \) is a critical pt, \( f_x(x_0) f_y(x_0) = 0. \)

If \( f_{yx} = f_{xy}, \) matrix \( A \) is real symmetric \((\therefore \text{real eigenvalues})\).
Examples:

i) \((15.2. \circ) \ f(x, y) = xy \ln(x^2 + y^2)\) Calculate \(\frac{\partial f}{\partial x}(y, y)\).

\[
\frac{\partial f}{\partial x} = \ln(x^2 + y^2) + xy \cdot \frac{2x}{x^2 + y^2}
\]

\[
\frac{\partial f}{\partial x} = \ln(y^2) + y \cdot \frac{2y}{y^2} = \ln(y^2) - \frac{y^2}{y^2} - y
\]

ii) \((15.4.b) \ f(x, y) = x^2y^5 - 4x^4y^3\) Calculate \(\frac{\partial^2 f}{\partial y \partial x}\)

\[
\frac{\partial f}{\partial x} = 2xy^5 - 24x^5y^3 \quad \frac{\partial^2 f}{\partial y \partial x} = 10xy^4 - 24x^5y^3
\]

Chain Rules:

i) \(z(t) = y(x(t))\) where \(y: \mathbb{R} \to \mathbb{R}\) \(\times: \mathbb{R} \to \mathbb{R}\) are both \(C^4\) functions.

\[
\frac{dz}{dt} = \lim_{\Delta t \to 0} \left( \frac{y(x(t+\Delta t)) - y(x(t))}{\Delta t} \right) = \lim_{\Delta t \to 0} \left( y'(x^*) \cdot \frac{(x(t+\Delta t) - x(t))}{\Delta t} \right)
\]

(where \(x^*\) lies between \(x(t+\Delta t)\) \& \(x(t)\))

\[
= y'(x(t)) \cdot x'(t) \quad \text{since} \quad x^* \to x(t) \text{ as } \Delta t \to 0.
\]

ii) \(z(t) = w(x(t), y(t))\). Assuming \(w(x, y), x(t), y(t)\) have the requisite smoothness:

\[
\frac{dz}{dt} = \frac{\partial w}{\partial x}(x(t), y(t)) \frac{dx(t)}{dt} + \frac{\partial w}{\partial y}(x(t), y(t)) \frac{dy(t)}{dt}
\]

Remark: We need to justify:

\[
\frac{dz}{dt} = \lim_{\Delta t \to 0} \left( \frac{w(x(t+\Delta t), y(t+\Delta t)) - w(x(t), y(t))}{\Delta t} \right)
\]

\[
= \lim_{\Delta t \to 0} \left( \frac{w(x(t+\Delta t), y(t+\Delta t)) - w(x(t), y(t+\Delta t))}{\Delta t} + \frac{w(x(t), y(t+\Delta t)) - w(x(t), y(t))}{\Delta t} \right)
\]

\[
= \lim_{\Delta t \to 0} \left( \frac{\partial w}{\partial x}(x^*, y(t+\Delta t))(x(t+\Delta t) - x(t)) + \frac{\partial w}{\partial y}(x(t), y^*)\left( y(t+\Delta t) - y(t) \right) \right)
\]

\[
= \frac{\partial w}{\partial x} \frac{dx(t)}{dt} + \frac{\partial w}{\partial y} \frac{dy(t)}{dt}
\]

Example: \(w(x, y) = x^2y + \sin(xy^2)\)

\(x(t) = t, \quad y(t) = t^2\)
a) Directly: \[ z(t) = t^4 + \sin(t^5) \quad \frac{dz}{dt} = 4t^3 + \cos(t^5) \cdot 5t^4 \]

b) Chain Rule: \[ \frac{dw}{dx} = 2xy + \cos(xy^2) \cdot y^2 \quad \frac{dx}{dt} = 1 \]
\[ \frac{dw}{dy} = x^2 + \cos(xy^2) \cdot 2xy \quad \frac{dy}{dt} = 2t \]
\[ \frac{dz}{dt} = \left( 2t^2 + t^4 \cos(t^5) \right) \cdot 1 + \left( t^2 + \cos(t^5) \cdot 2t^3 \right) \cdot 2t \]
\[ = 4t^3 + 5t^4 \cos(t^5) \]

(iii) General case: \[ w = w(\mathbf{x}), \quad \mathbf{x} = (x_1, x_2, \ldots, x_n) \]
\[ z(t) = w(\mathbf{x}(t)) = w(x_1(t), x_2(t), \ldots, x_n(t)) \]
\[ \frac{dz}{dt} = \sum_{i=1}^{n} \frac{\partial w}{\partial x_i} \frac{dx_i}{dt} \quad (= \nabla w(\mathbf{x}(t)) \cdot \mathbf{x}'(t), \text{where} \quad \nabla w = (w_{x_1}, w_{x_2}, \ldots, w_{x_n})) \]

Chain Rule with Partial Derivatives:

i) \[ z(u, v) = y(x(u, v)) \]
\[ \frac{dz}{du} = y'(x(u, v)) \frac{dx}{du} \quad \frac{dz}{dv} = y'(x(u, v)) \frac{dx}{dv} \]

ii) \[ z(u, v) = W(x(u, v), y(u, v)) \]
\[ \frac{dz}{du} = \frac{\partial w}{\partial x}(x(u, v), y(u, v)) \frac{dx}{du} + \frac{\partial w}{\partial y}(x(u, v), y(u, v)) \frac{dy}{du} \]
Similarly, for \[ \frac{dz}{dv} \]

Derivatives of \( f : \mathbb{R}^n \rightarrow \mathbb{R}^m \): (Ch.16, p.112)

\[ f(\mathbf{x}) = \begin{bmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{bmatrix} \]

Heuristics: If \( m = n = 1 \), \( f(\mathbf{x}) = f(x_0) + f'(x_0)(\mathbf{x} - \mathbf{x}_0) + \ldots \)
and \( f'(x_0) \) is the coefficient of the linear term in the Taylor expansion.
Consider the case \( m = 2, n = 3 \):
\[
F(x) = \begin{bmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \end{bmatrix}
\]

We anticipate Taylor series and assume that in some neighborhood of \( x_0 = (x_{01}, x_{02}, x_{03}) \):
\[
f_1(x) = f_1(x_0) + \sum_{j=1}^{3} \frac{\partial f_1}{\partial x_j} (x_j - x_{0j}) + \cdots
\]
\[
f_2(x) = f_2(x_0) + \sum_{j=1}^{3} \frac{\partial f_2}{\partial x_j} (x_j - x_{0j}) + \cdots
\]
\[
\therefore F(x) = F(x_0) + \left[ \begin{array}{ccc}
\frac{\partial f_1}{\partial x_1} (x_1 - x_{01}) & \frac{\partial f_1}{\partial x_2} (x_2 - x_{02}) & \frac{\partial f_1}{\partial x_3} (x_3 - x_{03}) \\
\frac{\partial f_2}{\partial x_1} (x_1 - x_{01}) & \frac{\partial f_2}{\partial x_2} (x_2 - x_{02}) & \frac{\partial f_2}{\partial x_3} (x_3 - x_{03}) \\
\end{array} \right] + \cdots
\]
\[
DF(x)
\]

The derivative of \( F \) at \( x_0 \) will be the \( 2 \times 3 \) (generally \( mxn \)) matrix of partial derivatives, \( DF(x_0) \), \( (total \ derivative \ matrix) \).

**Total Derivative Matrix:** (p113-114).

\( f: \mathbb{R}^n \rightarrow \mathbb{R}^m \)

The linear approximation to \( f \) at \( x_0 \) has the form:
\[
F(x) \approx F(x_0) + DF(x_0) (x - x_0) \quad \text{where} \quad \lim_{x \to x_0} \frac{F(x) - F(x_0) - DF(x_0)}{x-x_0} = 0
\]

We say that \( f \) is differentiable at \( x_0 \) and the total derivative matrix
\[
DF(x_0) = \begin{bmatrix}
\frac{\partial f_1}{\partial x_1} (x_1 - x_{01}) & \frac{\partial f_1}{\partial x_2} (x_2 - x_{02}) & \cdots & \frac{\partial f_1}{\partial x_n} (x_n - x_{0n}) \\
\frac{\partial f_2}{\partial x_1} (x_1 - x_{01}) & \frac{\partial f_2}{\partial x_2} (x_2 - x_{02}) & \cdots & \frac{\partial f_2}{\partial x_n} (x_n - x_{0n}) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial f_m}{\partial x_1} (x_1 - x_{01}) & \frac{\partial f_m}{\partial x_2} (x_2 - x_{02}) & \cdots & \frac{\partial f_m}{\partial x_n} (x_n - x_{0n}) \\
\end{bmatrix} \quad (mxn)
\]

**Example:** \( F(x, y) = \begin{bmatrix} x^2y^4 \\ \cos(\pi(x^2+y^2)) \\ e^{x-y} \end{bmatrix} \)

Compute \( DF(x) \) and the linear approximation at \( x_0 = (\frac{\pi}{2}, \frac{\pi}{2}) \).
\[ \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 4y^3 \\ -\sin(m\omega_1\omega_2)(2\pi x) & -\sin(m\omega_1\omega_2)(2\pi y) \\ e^{-y} & -e^{-y} \end{bmatrix} \]

\[ \begin{bmatrix} \frac{\partial f}{\partial x} \bigg|_{(\frac{\pi}{4}, \frac{\pi}{4})} = \begin{bmatrix} 1 \\ -\pi \\ 1 \end{bmatrix} \]

\[ \begin{bmatrix} \frac{\partial f}{\partial y} \bigg|_{(\frac{\pi}{4}, \frac{\pi}{4})} = \begin{bmatrix} \frac{\pi}{16} \\ 0 \\ 1 \end{bmatrix} \]

\[ \left[ \begin{bmatrix} x^2 + y^2 \\ \cos(m\omega_1\omega_2) \\ e^{-y} \end{bmatrix} \right] \approx \begin{bmatrix} \frac{\pi}{16} \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 & \frac{\pi}{2} \\ -\pi & -\pi \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x - \frac{\pi}{2} \\ y - \frac{\pi}{2} \end{bmatrix} \]

**Remark:** Recall that when \( m=1, n=2 \), i.e., \( f: \mathbb{R}^2 \rightarrow \mathbb{R} \) and \( z = f(x,y) \), the linear approximation:

\[ z \approx f(x_0, y_0) + \left. \frac{\partial f}{\partial x}(x_0, y_0) \right|_{x=x_0, y=y_0} (x-x_0) + \left. \frac{\partial f}{\partial y}(x_0, y_0) \right|_{x=x_0, y=y_0} (y-y_0) \]

defines the plane tangent to the surface \( z = f(x,y) \) at the point \( (x_0, y_0, z_0) \).

**Theorem 1.6.9**: Let \( f: \mathbb{R} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m \) and let \( \mathbf{x}_0 \in \mathbb{R}^n \). Then:

\[ f \text{ differentiable at } \mathbf{x}_0 \Rightarrow f \text{ continuous at } \mathbf{x}_0 \]

**Theorem 1.6.10**: Let \( f: \mathbb{R} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m \). If the partial derivatives of \( f \) exist and are continuous in \( B_\varepsilon(\mathbf{x}_0) = \{ \mathbf{x} : ||\mathbf{x} - \mathbf{x}_0|| < \varepsilon \} \subseteq \mathbb{R}^n \) then \( f \) is differentiable at \( \mathbf{x}_0 \).

**Note:**

1) \( f \) differentiable at \( \mathbf{x}_0 \) \( \Rightarrow \) partial derivatives exist at \( \mathbf{x}_0 \).

2) all partial derivatives exist \( \Rightarrow \) are continuous in some neighborhood of \( \mathbf{x}_0 \)

\( \Rightarrow f \) differentiable at \( \mathbf{x}_0 \).
Example:
Let \( \varrho(x) = \begin{cases} \varrho, & x = 0 \\ \chi^2 \sin \left( \frac{x}{\chi} \right) & \end{cases} \)
and let \( f(x,y) = \varrho(x) \varrho(y) \).

i) \( f(x,y) \) is continuous at \((0,0)\) since \( f(0,0) = 0 \) and \( |f(x,y) - 0| \leq r^{2g} \) which goes to 0 as \( r = \sqrt{x^2 + y^2} \to 0 \).

ii) Since \( f(x,0) = f(0,y) = 0 \), \( \frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0 \).

iii) Since \( f(0,0) = 0 \), the linear approximation would have the form
\[
f(x,y) \approx \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}
\]

iv) \( f \) is not differentiable at \((0,0)\) since:
\[
\lim_{x \to 0} \frac{(xy)^{1/3} \sin \left( \frac{y}{x} \right) \sin \left( \frac{x}{y} \right) - a_{11}x - a_{12}y}{\sqrt{x^2 + y^2}} = \lim_{x \to 0} \frac{r^{2g} \left( \cos \Theta \sin \Theta \right) \sin \left( \frac{r \cos \Theta}{r \sin \Theta} \right) \sin \left( \frac{r \sin \Theta}{r \cos \Theta} \right)}{r}
\]
does not exist for any \( a_{11}, a_{12} \).

v) Note that \( \frac{\partial f}{\partial x} \) and \( \frac{\partial f}{\partial y} \) are not continuous in any \( B_{\epsilon}(0) \).
\[
\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} \varrho(x) \varrho(y) = \frac{1}{3} \chi^{-2g} \sin \left( \frac{x}{\chi} \right) + \chi^{1/3} \cos \left( \frac{x}{\chi} \right) \left( -\frac{2}{\chi^2} \right) \varrho(y)
\]
\[
= \frac{1}{3} \left[ r^{-1/3} \cos \Theta \right]^{1/3} \sin \left( \frac{r \cos \Theta}{r \sin \Theta} \right) - 2 r^{-4/3} \cos \Theta \sin \left( \frac{r \cos \Theta}{r \sin \Theta} \right) \right] \sin \left( \frac{r \sin \Theta}{r \cos \Theta} \right) \sin \left( \frac{r \sin \Theta}{r \cos \Theta} \right)
\]

For example, fixing \( \Theta = \frac{7\pi}{4} \) and letting \( r \to 0 \):
\[
\lim_{r \to 0} f_x = \frac{1}{3} \left[ r^{-1/3} \cdot 2^{1/3} \sin \left( \frac{\sqrt{2}}{r} \right) - 2 r^{-4/3} \cdot 2^{1/3} \cos \left( \frac{\sqrt{2}}{r} \right) \right] 2^{1/3} \sin \left( \frac{\sqrt{2}}{r} \right)
\]
does not exist.
Gradient, Divergence & Curl: (p.122)

**Def:** Let \( f: \mathbb{R}^n \to \mathbb{R} \). If \( f \) is differentiable we define:

\[
\nabla f: \mathbb{R}^n \to \mathbb{R}^n \text{ as } \nabla f = \begin{bmatrix}
\frac{\partial f}{\partial x_1} \\
\frac{\partial f}{\partial x_2} \\
\vdots \\
\frac{\partial f}{\partial x_n}
\end{bmatrix} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} e_i \quad (e_i = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}^T \text{ i^{th} row}).
\]

**Notation:** \( \text{grad } f \) also used.

**Remark:** We will interpret \( \nabla f \) geometrically as a vector. In particular, if \( n=3 \), we will have:

\[
\nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}
\]

**Directional Derivative:**

As an example, consider the graph of \( f: \mathbb{R} \to \mathbb{R}^2 \). The graph \( z = f(x,y) \) defines a surface. At a pt \( (x_0,y_0) \in \mathbb{R} \), the partial derivatives \( \frac{\partial f}{\partial x}(x_0,y_0) \) and \( \frac{\partial f}{\partial y}(x_0,y_0) \) compute rates of change in the respective coordinate directions. What about other directions?

**General Case:** \( f: \mathbb{R}^n \to \mathbb{R}, \quad x_0 \in \mathbb{R} \)

Let \( \mathbf{u} = (u_1, u_2, \ldots, u_n) \) be an \( n \)-dimensional unit vector, ie \( |\mathbf{u}| = \sum_{i=1}^{n} u_i^2 = 1 \). A line in \( \mathbb{R}^n \) passing through \( x_0 \) in the \( \mathbf{u} \)-direction is parametrically defined as: \( \mathbf{x}(t) = x_0 + t \mathbf{u} \) (Note that \( t \) has the dimensions of \( \mathbb{R} \)).

The directional derivative is given by:

\[
\left. \frac{d}{dt} f(x_0 + t \mathbf{u}) \right|_{t=0} = \left. \frac{d}{dt} f(x_0 + t u_1, \ldots, x_0 + t u_n) \right|_{t=0} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0,\ldots,x_0) \cdot u_i
\]

\[
= \nabla f(x_0) \cdot \mathbf{u} \quad \text{(Recall the Chain Rule)}
\]
Example: The elevation of a mound is given by \( f(x, y) = \frac{5}{1 + 4x^2 + y^2} \), where distance is measured in hundreds of feet and the positive \( x \)-axis points due east. Determine the rate of descent at a point on the hill \((x_0, y_0, z_0) = (1, 2, 5\frac{2}{3})\) in the direction 30° north of east.

\[
\frac{df}{dx} = \frac{-5}{(1 + 4x^2 + y^2)^2} \cdot 8x = \frac{-40x}{(1 + 4x^2 + y^2)^2} \quad \frac{df}{dy} = \frac{-10y}{(1 + 4x^2 + y^2)^2}
\]

\[
\frac{df}{dx}(1, 2) = -40 \quad \frac{df}{dy}(1, 2) = -20
\]

\[
y = \cos 30^\circ i + \sin 30^\circ j = \frac{\sqrt{3}}{2} i + \frac{1}{2} j
\]

\[
\mathbf{D}_x f(1, 2) = \nabla f(1, 2) \cdot \mathbf{y} = \left( -\frac{40}{81} \mathbf{i} + \frac{20}{81} \mathbf{j} \right) \cdot \left( \frac{\sqrt{3}}{2} \mathbf{i} + \frac{1}{2} \mathbf{j} \right) = \frac{-20\sqrt{3} - 10}{81} = -0.5511 \text{ ft}
\]

Example: A thermal source, located at the origin of a coordinate system, maintains a steady state temperature in a region described by:

\[
T(x, y, z) = 90 \exp \left\{ -\frac{x^2}{4} - \frac{y^2}{9} - \frac{z^2}{16} \right\} \text{ °F}, \text{ where distance is measured in ft.}
\]

Determine the temperature gradient (i.e., the rate of change of temperature with respect to distance) at location \((x_0, y_0, z_0) = (1, 2, 3)\) in the direction defined by \(\mathbf{z} = \frac{1}{2} \mathbf{i} + 2 \mathbf{j} - 2 \mathbf{k}\).

\[
\nabla T = T_x \mathbf{i} + T_y \mathbf{j} + T_z \mathbf{k} = 90e^{-\frac{x^2}{4}-\frac{y^2}{9}-\frac{z^2}{16}} \left( -\frac{x}{2} \mathbf{i} - \frac{2y}{9} \mathbf{j} - \frac{z}{8} \mathbf{k} \right)
\]

\[
\nabla T(1, 2, 3) = 90e^{-\frac{1}{4}-\frac{4}{9}-\frac{9}{16}} \left( -\frac{1}{2} \mathbf{i} - \frac{4}{9} \mathbf{j} - \frac{3}{8} \mathbf{k} \right)
\]

Need unit vector: \(\mathbf{y} = \frac{1}{\sqrt{69}} \mathbf{x} = \frac{1}{2} \mathbf{i} + \frac{2}{3} \mathbf{j} - \frac{2}{3} \mathbf{k}
\]

\[
\mathbf{D}_x T(1, 2, 3) = 90e^{-\frac{1}{4}-\frac{4}{9}+\frac{1}{4}} = 90e^{-\frac{18\sqrt{144}}{108}} \text{ °F/ft}
\]

Remarks:

i) Partial derivatives \(\frac{df}{dx}, \frac{df}{dy}\), etc., correspond to directional derivatives in the coordinate axis directions (e.g., \(\frac{df}{dx} = \mathbf{D}_x f = \nabla f \cdot \mathbf{i}\), etc.)

ii) Since \(\mathbf{D}_x f = \nabla f \cdot \mathbf{y} = |\nabla f| |\mathbf{y}| \cos \theta = |\nabla f| \cos \theta\), where \(\theta\) is the angle between \(\nabla f \mathbf{y}\), the maximum value of the directional derivative at a point occurs in the direction of \(\nabla f\) (when \(\theta = 0\)). The minimum value occurs in the direction of \(-\nabla f\) (when \(\theta = \pi\)).
iii) Recall that if $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x,y) = c$ defines a level curve, a contour line whose points on the surface $z = f(x,y)$ have height $c$ above the plane $z = 0$.

If $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x,y,z) = c$ defines a level surface.

Claim:

1) If $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ is differentiable, (and $\nabla f(x_o) \neq 0$ at some $x_o \in \mathbb{R}^2$), then $\nabla f(x_o)$ is a vector in $\mathbb{R}^2$ that is perpendicular to the level curve $f(x) = f(x_o) = c$ at $x_o$.

Proof:

Let $z(t)$ parameterize the level curve $f(x) = f(x_o)$, i.e., $f(z(t)) = f(x_o)$. Let $z(t_o) = x_o$. Then:

$$\frac{d}{dt} f(z(t)) \bigg|_{t=t_o} = \nabla f(z(t), z'(t)) \bigg|_{t=t_o} = \nabla f(x_o) \cdot z'(t_o) = 0$$

Recall that $z'(t_o)$ is tangent to the level curve at $x_o$; $\nabla f(x_o)$ is perpendicular to the level curve at $x_o$.

Example: $f(x,y) = 4x^2 + y^2$. Let $x_o = (x_o, y_o) = (1, 2)$. The level curve passing through $(1,2)$ is the ellipse $4x^2 + y^2 = 8$, $\left(\frac{x^2}{2} + \frac{y^2}{8} = 1\right)$.

$\nabla f = 8x \mathbf{i} + 2y \mathbf{j}$. $\nabla f(1,2) = 8 \mathbf{i} + 4 \mathbf{j}$.
(i) Let $f: \mathbb{R}^2 \to \mathbb{R}$ be differentiable and $\nabla f(\chi_0) \neq \mathbf{0}$ at some $\chi_0 \in S$. Then $\nabla f(\chi_0)$ is $\perp$ to the level surface $f(\chi) = f(\chi_0) = c$ at point $\chi_0$.

**Proof:**
Let $\chi_1(t)$ and $\chi_2(s)$ be two arbitrary curves on the level surface $f(\chi) = c (= f(\chi_0))$ passing through pt $\chi_0$. Let $\chi_1(t_0) = \chi_2(s_0) = \chi_0$. Then $\nabla f(\chi_0)$ is $\perp$ to both $\chi'_1(t_0)$ and $\chi'_2(s_0)$. \(\therefore\) $\nabla f(\chi_0)$ is $\perp$ to the plane tangent to the surface $f(\chi) = c$ at pt $\chi_0$. (See below)

**Example:** Let $f(\chi) = f(x, y, z) = x^2 + 4y^2 + z^2$ \& $\chi_0 = (1, 1, 2)$. The surface is the ellipsoid of revolution $x^2 + 4y^2 + z^2 = 9$ or $\frac{x^2}{9} + \frac{y^2}{\frac{9}{4}} + \frac{z^2}{9} = 1$. $\nabla f(\chi) = 2x_i + 8y_j + 2z_k$. $\nabla f(\chi_0) = 2\mathbf{i} + 8\mathbf{j} + 4\mathbf{k}$.

Figure for proof:

**Mean Value Theorem:** Let $f: S \subseteq \mathbb{R}^n \to \mathbb{R}$ be $C^1(S)$ and let $S$ be convex (i.e., if $\chi_1, \chi_2 \in S$, then the line segment connecting them lies in $S$). Then $\exists$ $\tilde{\chi} \in S$ such that:

$$f(\chi_2) - f(\chi_1) = \nabla f(\tilde{\chi}) \cdot (\chi_2 - \chi_1).$$

**Proof:** Let $\gamma_2 + (1-t)\chi_1$, $0 \leq t \leq 1$ parameterize the line segment in $S$ connecting $\chi_1$ \& $\chi_2$. Consider \(g(t) = f(\gamma_2 + (1-t)\chi_1), 0 \leq t \leq 1\) and apply the Mean Value Theorem.