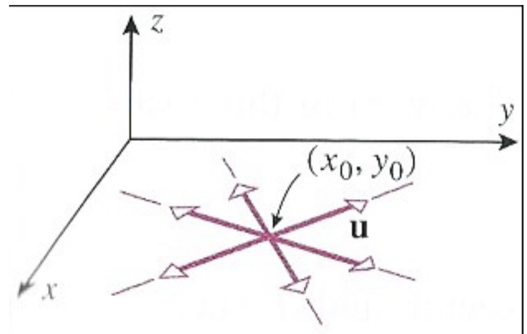
7 Directional Derivatives and Gradients

Suppose we need to compute the rate of change of $f(x, y)$ with respect to the distance from a point (a, b) in some direction. Let $\vec{u} = u_1\vec{i} + u_2\vec{j}$ be the unit vector that has its initial point at (a, b) and points in the desired direction. It determines a line in the xy-plane:



$$x = a + s\,u_1\,, \quad y = b + s\,u_2$$

where s is the arc length parameter that has its reference point at (a, b) and has positive values in the direction of $\vec{u}$.

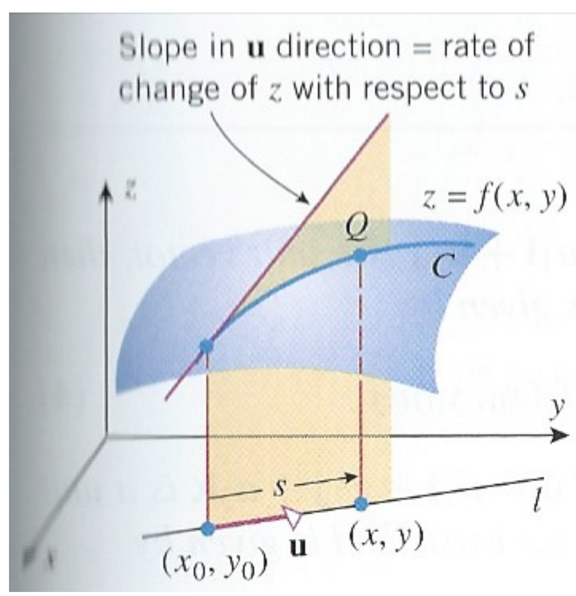
Definition. The **directional derivative** of $f(x, y)$ in the direction of $\vec{u}$ at (a, b) is denoted by $D_{\vec{u}}f(a, b)$ and is defined by

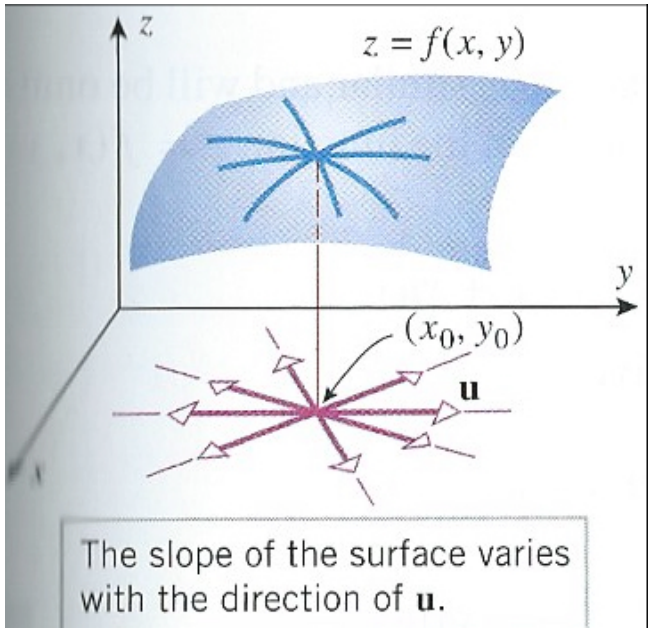
$$D_{\vec{u}}f(a, b) = \frac{d}{ds}\left[f(a + s\,u_1\,,\, b + s\,u_2)\right]\Big|_{s=0} = f_x(a, b)\,u_1 + f_y(a, b)\,u_2$$

provided this derivative exists.

Analytically, $D_{\vec{u}}f(a, b)$ is the instantaneous rate of change of $f(x, y)$ with respect to the distance in the direction of $\vec{u}$ at the point (a, b).

Geometrically, $D_{\vec{u}}f(a, b)$ is the slope of the surface $z = f(x, y)$ in the direction of $\vec{u}$ at the point $(a, b, f(a, b))$.





Generalisation to $f(x, y, z)$ (and $f(x_1, \ldots, x_n)$) is straightforward.

Definition. Let $\vec{u} = u_1 \vec{i} + u_2 \vec{j} + u_3 \vec{k}$ be a unit vector.
The **directional derivative** of $f(x, y, z)$ in the direction of $\vec{u}$ at (a, b, c) is denoted by $D_{\vec{u}} f(a, b, c)$ and is defined by

$$\begin{aligned} D_{\vec{u}} f(a, b, c) &= \frac{d}{ds} \left[f(a + s\,u_1\,,\ b + s\,u_2\,,\ c + s\,u_3) \right] \Big|_{s=0} \\ &= f_x(a, b, c)\,u_1 + f_y(a, b, c)\,u_2 + f_z(a, b, c)\,u_3 \end{aligned}$$

Example. Find $D_{\vec{u}} f(2, 1)$ in the direction of $\vec{a} = 3\,\vec{i} + 4\,\vec{j}$

$$f(x, y) = \ln \left(\frac{1}{2} e^{2/3} \sqrt[3]{12 \sin(x - 2y) + 8y^2 - x^3 - 6x^2 y + 32} \right)$$

Answer: $D_{\vec{u}} f(2, 1) = -5/3$

The gradient

Note that

$$D_{\vec{u}}f = f_x\,u_1 + f_y\,u_2 + f_z\,u_3 = (f_x\,\vec{i} + f_y\,\vec{j} + f_z\,\vec{k}) \cdot (u_1\,\vec{i} + u_2\,\vec{j} + u_3\,\vec{k})$$

Definition. Let $\vec{e}_i$ be the standard orthonormal coordinate basis of $\mathbb{R}^n$, so that $\vec{r} = \sum_{i=1}^n x_i\vec{e}_i$.
The **gradient** of $f(x_1, \cdots, x_n)$ is defined by

$$\vec{\nabla} f(x_1, \cdots, x_n) = \sum_{i=1}^{n} \frac{\partial f(x_1, \cdots, x_n)}{\partial x_i}\vec{e}_i$$

In particular

$$\vec{\nabla} f(x, y) = f_x(x, y)\,\vec{i} + f_y(x, y)\,\vec{j}$$

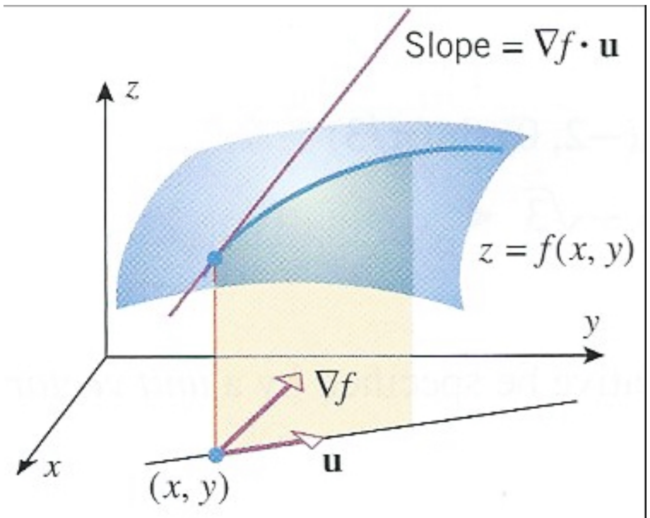
$$\vec{\nabla} f(x, y, z) = f_x(x, y, z)\,\vec{i} + f_y(x, y, z)\,\vec{j} + f_z(x, y, z)\,\vec{k}$$

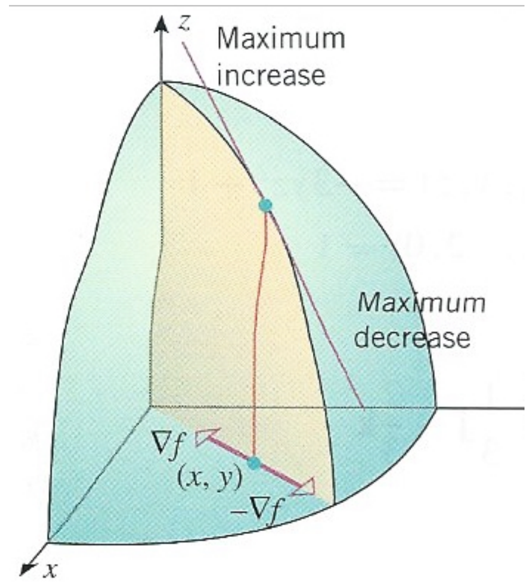
The symbol $\vec{\nabla}$ is read as either "nabla" (from ancient Hebrew) or "del" (it is inverted Δ).

$$D_{\vec{u}}f(a, b) = \vec{\nabla} f(a, b)\cdot\vec{u}\,, \quad D_{\vec{u}}f(a, b, c) = \vec{\nabla} f(a, b, c)\cdot\vec{u}\,, \quad D_{\vec{u}}f = \vec{\nabla} f\cdot\vec{u}$$

Example. Find $\vec{\nabla} r$; $r = \sqrt{x^2 + y^2 + z^2}$ and $D_{\vec{u}}r(1, 1, 1)$ in the direction of $\vec{a} = \vec{i} + 2\,\vec{j} + 2\,\vec{k}$.

Properties of the gradient





$$D_{\vec{u}}f(a,b) = \vec{\nabla}f(a,b) \cdot \vec{u} = |\vec{\nabla}f(a,b)|\,|\vec{u}|\,\cos\theta = |\vec{\nabla}f(a,b)|\,\cos\theta$$

Since $-1 \leq \cos\theta \leq 1$, if $|\vec{\nabla}f(a,b)| \neq 0$ then the maximum value of $D_{\vec{u}}f(a,b)$ is $|\vec{\nabla}f(a,b)|$ and it occurs when $\theta = 0$, that is, when $\vec{u}$ is in the direction of $\vec{\nabla}f(a,b)$.

Geometrically, the maximum slope of the surface $z = f(x,y)$ at (a,b) is in the direction of the gradient and is equal to $|\vec{\nabla}f(a,b)|$.

If $|\vec{\nabla}f(a,b)| = 0$ then $D_{\vec{u}}f(a,b) = 0$ in all directions at (a,b).
It occurs where the surface $z = f(x,y)$ has a relative maximum or minimum or a saddle point.

Since $D_{\vec{u}}f(x_1, \ldots, x_n) = |\vec{\nabla} f(x_1, \ldots, x_n)| \cos\theta$, these properties hold for functions of any number of variables.

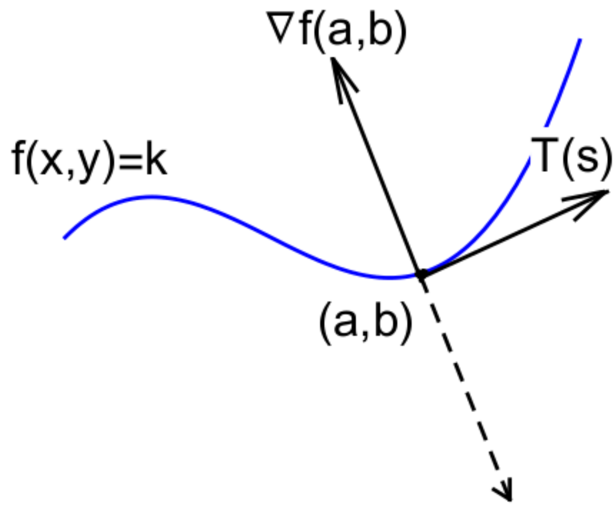
Theorem. Let f be a function differentiable at a point P.

1. If $\vec{\nabla} f = \vec{0}$ at P then all directional derivatives of f at P are 0.
2. If $\vec{\nabla} f \neq \vec{0}$ at P then the derivative in the direction of $\vec{\nabla} f$ at P has the largest value equal to $|\vec{\nabla} f|$ at P.
3. If $\vec{\nabla} f \neq \vec{0}$ at P then the derivative in the direction opposite to that of $\vec{\nabla} f$ at P has the smallest value equal to $-|\vec{\nabla} f|$ at P.

Example. The point $P = (2, 3, -1)$

$$f(x, y, z) = \sqrt{2xy + 3z^4 - 6\cos(3x - 2y)}$$

Gradients are normal to level curves and level surfaces



Level curve C: $f(x, y) = k$.

Let C be smoothly parametrised as $x = x(s)$, $y = y(s)$ where s is an arc length parameter. The unit tangent vector to C is

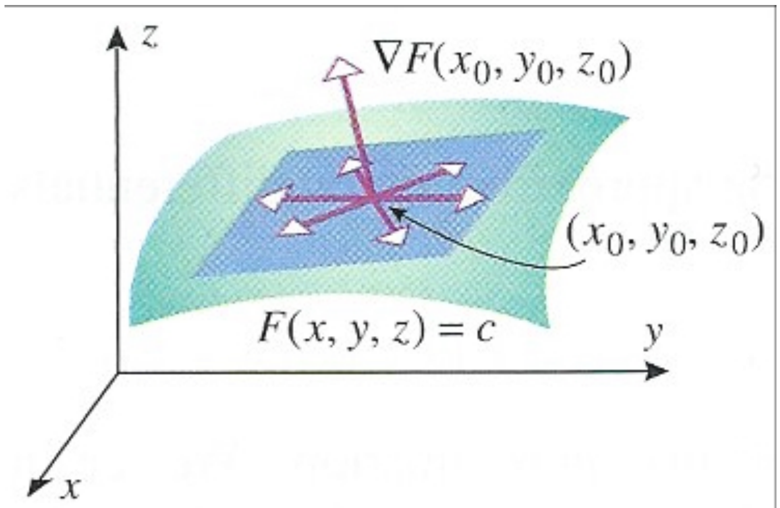
$$\vec{T}(s) = \frac{dx}{ds}\,\vec{i} + \frac{dy}{ds}\,\vec{j}$$

Since $f(x, y)$ is constant on C we expect $D_{\vec{T}} f(x, y) = 0$. Indeed

$$\begin{aligned} D_{\vec{T}} f(x, y) &= \vec{\nabla} f \cdot \vec{T} = (f_x\,\vec{i} + f_y\,\vec{j}) \cdot (\frac{dx}{ds}\,\vec{i} + \frac{dy}{ds}\,\vec{j}) \\ &= f_x \frac{dx}{ds} + f_y \frac{dy}{ds} = \frac{d}{ds} f(x(s), y(s)) = 0 \;\Rightarrow\; \vec{\nabla} f \perp \vec{T} \end{aligned}$$

Thus if (a, b) belongs to the level curve, and $\vec{\nabla} f(a, b) \neq \vec{0}$ then $\vec{\nabla} f(a, b)$ is normal to $\vec{T}$ at (a, b) and therefore to the level curve.

Definition. A vector is called normal to a surface at (a, b, c) if it is normal to a tangent vector to any curve on the surface through (a, b, c).



Level surface σ: $F(x, y, z) = k$

Let C, smoothly parametrised as $x = x(s)$, $y = y(s)$, $z = z(s)$

be any curve on σ through (a, b, c). The unit tangent vector to C is

$$\vec{T}(s) = \frac{dx}{ds}\,\vec{i} + \frac{dy}{ds}\,\vec{j} + \frac{dz}{ds}\,\vec{k}$$

and $D_{\vec{T}}F(x, y, z)$ is

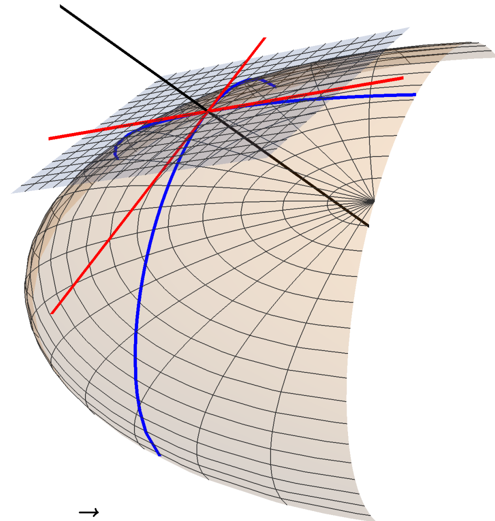
$$D_{\vec{T}}F(x, y, z) = \vec{\nabla}F \cdot \vec{T} = (F_x\,\vec{i} + F_y\,\vec{j} + F_z\,\vec{k}) \cdot (\frac{dx}{ds}\,\vec{i} + \frac{dy}{ds}\,\vec{j} + \frac{dz}{ds}\,\vec{k})$$

$$= F_x\frac{dx}{ds} + F_y\frac{dy}{ds} + F_z\frac{dz}{ds} = \frac{d}{ds}F(x(s), y(s), z(s)) = 0 \;\Rightarrow\; \vec{\nabla}F \perp \vec{T}$$

Thus, $\vec{\nabla}F(a, b, c)$ is normal to $\vec{T}$ at (a, b, c) and therefore to σ.

Tangent planes

Consider a level surface σ: $F(x, y, z) = k$,
and let $P = (a, b, c)$ belong to σ.
Since $\vec{\nabla} F(a, b, c)$ is normal to tangent vectors to curves on σ through P,
all these tangent vectors belong to one and the same plane.
This plane is called the **tangent plane** to the surface σ at P.



To find an equation of the tangent plane we use that if we know a vector $\vec{n}$ normal to a plane through a point $\vec{r}_0 = a\,\vec{i} + b\,\vec{j} + c\,\vec{k}$ then an equation of the plane is

$$\vec{n} \cdot (\vec{r} - \vec{r}_0) = 0 \quad \Leftrightarrow \quad n_1(x-a) + n_2(y-b) + n_3(z-c) = 0$$

because $\vec{r} - \vec{r}_0$ is parallel to the plane and therefore normal to $\vec{n}$.

Choosing $\vec{n} = \vec{\nabla} F(a, b, c)$, we get the equation of the tangent plane to the level surface σ at $P = (a, b, c)$

$$F_x(a,b,c)(x-a) + F_y(a,b,c)(y-b) + F_z(a,b,c)(z-c) = 0$$

The line through P parallel to $\vec{\nabla} F(a, b, c)$ is perpendicular to the tangent plane, and is called the **normal line** to the surface σ at P. Its parametric equations are

$$x = a + F_x(a,b,c)t\,, \quad y = b + F_y(a,b,c)t\,, \quad z = c + F_z(a,b,c)t$$

Example. $4x^2 + y^2 + z^2 = 18$ at $(2, 1, 1)$.
Tangent plane, normal line, the angle the tangent plane makes with the xy-plane?

Tangent planes to $z = f(x, y)$

The graph of a function $z = f(x, y)$ can be thought of as the level surface of the function $F(x, y, z) = f(x, y) - z$ with constant 0.

We find

1. the gradient

$$\vec{\nabla} F(a, b, c) = f_x(a, b)\,\vec{i} + f_y(a, b)\,\vec{j} - \vec{k}\,, \quad c = f(a, b)$$

2. the equation of the tangent plane to the surface $z = f(x, y)$ at $(a, b, f(a, b))$

$$f_x(a, b)(x - a) + f_y(a, b)(y - b) - (z - c) = 0 \;\Rightarrow$$
$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

that is the local linear approximation of f at (a, b),

3. the parametric equations of the normal line to the surface $z = f(x, y)$ at $(a, b, f(a, b))$

$$x = a + f_x(a, b)\,t\,, \quad y = b + f_y(a, b)\,t\,, \quad z = f(a, b) - t$$

Example. Consider the surface

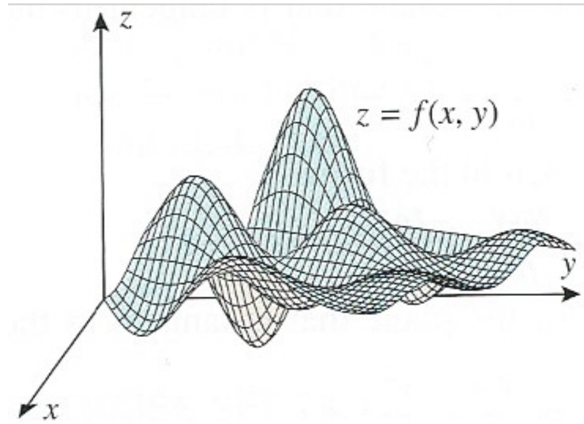
$$z = f(x, y) = \ln\left(\frac{1}{2} e^{2/3} \sqrt[3]{12\sin(x - 2y) + 8y^2 - x^3 - 6x^2 y + 32}\right)$$

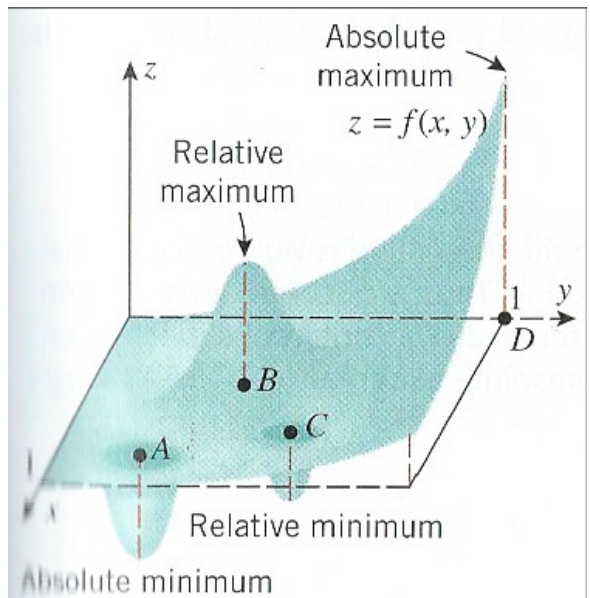
1. Find an equation for the tangent plane and parametric equations for the normal line to the surface at the point $P = (2, 1, z_0)$ where $z_0 = f(2, 1)$.

2. Find points of intersection of the tangent plane with the x-, y- and z-axes. Sketch the tangent plane, and show the point P on it. Sketch the normal line to the surface at P.

8 Maxima and minima of functions of two variables

Definition. A function f of two variables is said to have a **relative maximum (minimum)** at a point (a, b) if there is a disc centred at (a, b) such that $f(a,b) \geq f(x,y)$ $(f(a,b) \leq f(x,y))$ for all points (x, y) that lie inside the disc.

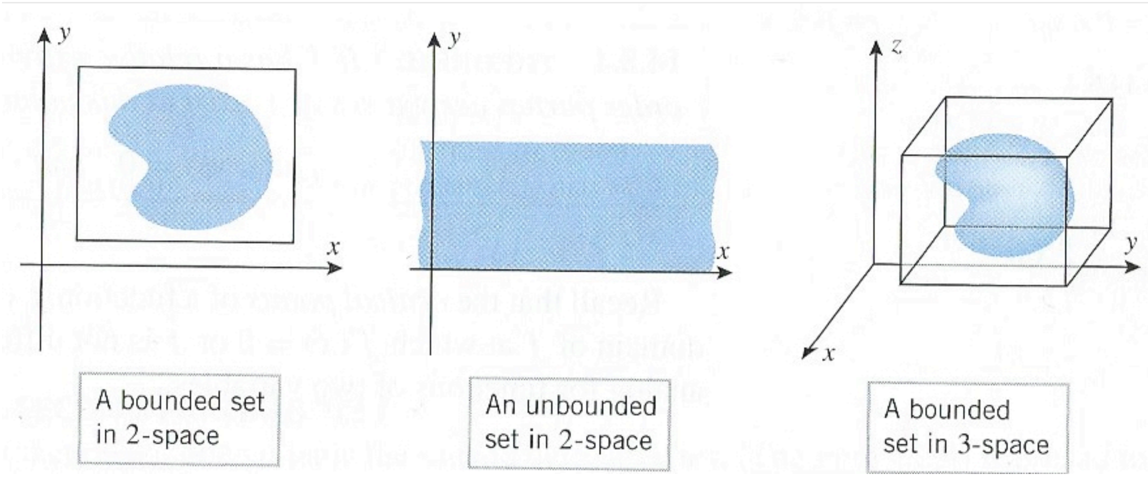




A function f is said to have an **absolute maximum (minimum)** at (a, b) if $f(a,b) \geq f(x,y)$ $(f(a,b) \leq f(x,y))$ for all points (x, y) that lie inside in the domain of f.

If f has a relative (absolute) maximum or minimum at (a, b) then we say that f has a **relative (absolute) extremum** at (a, b).

relative $\leftrightarrow$ local



The extreme-value theorem. If $f(x, y)$ is continuous on a closed and bounded set R, then f has both absolute maximum and an absolute minimum on R.

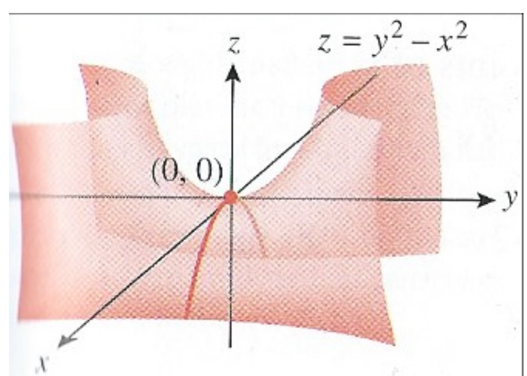
Finding relative extrema

Theorem. If f has a relative extremum at (a, b), and if the first-order derivatives of f exist at this point, then

$$f_x(a, b) = 0 \text{ and } f_y(a, b) = 0$$

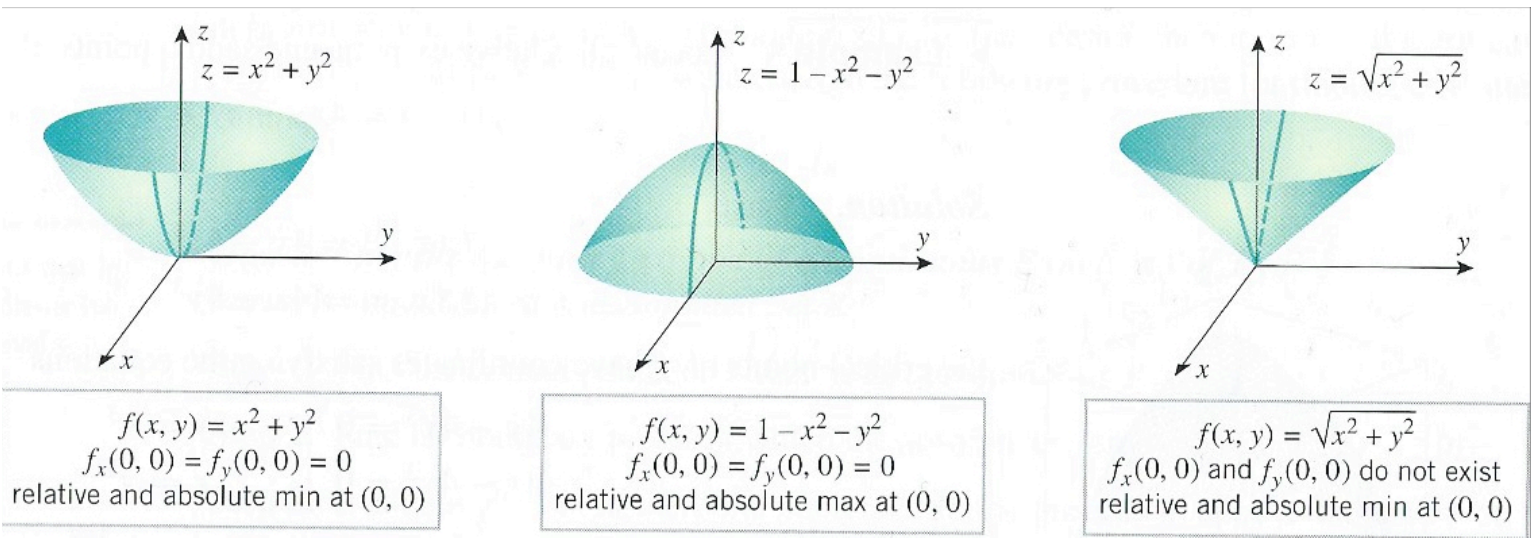
Definition. A point (a, b) in the domain of $f(x, y)$ is called a **critical point** of f if $f_x(a, b) = 0$ and $f_y(a, b) = 0$, or if one or both partial derivatives do not exist at (a, b).

Example. $f(x, y) = y^2 - x^2$ is a hyperbolic paraboloid.
$f_x = -2x$, $f_y = 2y \Rightarrow (0, 0)$ is critical but it is not a relative extremum.
It is a **saddle point**.



We say that a surface $z = f(x, y)$ has a **saddle point** at (a, b) if there are two distinct vertical planes through this point such that the trace of the surface in one of the planes has a relative maximum at (a, b), and the trace in the other has a relative minimum at (a, b).

Example.



How to determine whether a critical point is a max or min?

The second partials test

Theorem. Let $f(x,y)$ have continuous second-order partial derivatives in some disc centred at a critical point (a,b), and let

$$D = f_{xx}(a,b)f_{yy}(a,b) - \big(f_{xy}(a,b)\big)^2$$

1. If $D > 0$ and $f_{xx}(a,b) > 0$, then f has a relative minimum at (a,b).
2. If $D > 0$ and $f_{xx}(a,b) < 0$, then f has a relative maximum at (a,b).
3. If $D < 0$, then f has a saddle point at(a,b).
4. If $D = 0$, then no conclusion can be drawn.

Example.

$$f(x,y) = x^4 - x^2y + y^2 - 3y + 4$$

How to find the absolute extrema of a continuous function of two variables on a closed and bounded set R?

1. Find the critical points of f that lie in the interior of R.
2. Find all the boundary points at which the absolute extrema can occur.
3. Evaluate $f(x,y)$ at the found points. The largest of these values is the absolute maximum, and the smallest the absolute minimum.

Example.

$f(x,y) = 3x + 6y - 3xy - 7\,, \quad R$ is the triangle $(0,0), (0,3), (5,0)$

Lagrange multipliers

Extremum problems with constraints:

Find max or min of the function $f(x_1, \dots, x_n)$ subject to constraints

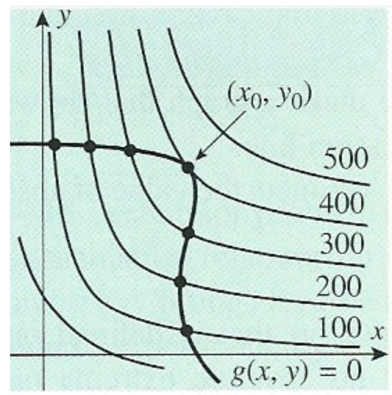
$g_\alpha(x_1, \dots, x_n)$, $\alpha = 1, \dots, m$

Consider $f(x, y)$ and $g(x, y) = 0$.

The graph of $g(x, y) = 0$ is a curve.

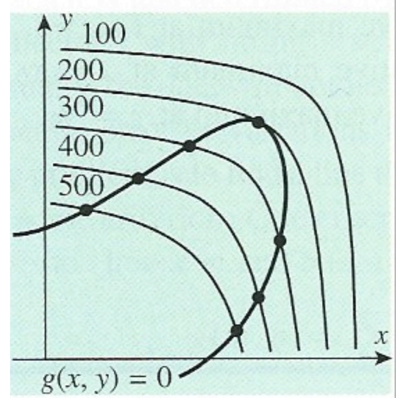
Consider level curves of f: $f(x, y) = k$.

At (a, b) the curves just touch, and thus have a common tangent line at (a, b). Since $\vec{\nabla} f(a, b)$ is normal to the level curve at (a, b), and $\vec{\nabla} g(a, b)$ is normal to the constraint curve at (a, b), we get $\vec{\nabla} f(a, b) || \vec{\nabla} g(a, b)$



Maximum of $f(x, y)$ is 400

(a)



Minimum of $f(x, y)$ is 200

$$\vec{\nabla} f(a, b) = \lambda \vec{\nabla} g(a, b)$$

for some scalar λ called the Lagrange multiplier.

Proof. Parametrise $g(x, y) = 0$.

Then, $f(x, y) = f(x(t), y(t))$ is a function of t and its local extrema are at

$$\frac{d}{dt} f(x(t), y(t)) = \frac{\partial f}{\partial x} x' + \frac{\partial f}{\partial y} y'$$
$$= \vec{\nabla} f \cdot (x' \vec{i} + y' \vec{j}) = \vec{\nabla} f \cdot \vec{T}$$

Thus, both $\vec{\nabla} f$ and $\vec{\nabla} g$ are $\perp$ to $\vec{T}$.

In general, we introduce a Lagrange multiplier λ_α for each of the constraint g_α, and the equations are

$$\vec{\nabla} f = \sum_{\alpha=1}^{m} \lambda_\alpha \, \vec{\nabla} g_\alpha \,.$$

Example. Find the points on the sphere $x^2 + y^2 + z^2 = 36$ that are closest to and farthest from the point $(1, 2, 2)$.