Queen Mary
University of London

MTH4101 Calculus II

Lecture notes for Week 3
Derivatives V

Thomas' Calculus, Sections 14.5 to 14.7

Rainer Klages

School of Mathematical Sciences
Queen Mary, University of London

Spring 2013

DEFINITION **Gradient Vector**

The **gradient vector (gradient)** of $f(x, y)$ at a point $P_0(x_0, y_0)$ is the vector

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

obtained by evaluating the partial derivatives of f at P_0.

The expression $\nabla f = \text{grad } f$ is called "grad f", "gradient of f", "del f" or "nabla f". We can now write the directional derivative using the gradient:

Theorem: Directional Derivative
If $f(x, y)$ is differentiable in an open region containing $P_0(x_0, y_0)$ then

$$\left(\frac{\mathrm{d}f}{\mathrm{d}s}\right)_{\mathbf{u},P_0} = (\nabla f)_{P_0} \cdot \mathbf{u}\,,$$

which is the scalar product of grad f at P_0 and $\mathbf{u}$.

Example:
Find the derivative of $f(x, y) = x\,e^y + \cos(xy)$ at the point $(2, 0)$ in the direction of $\mathbf{v} = 3\mathbf{i} - 4\mathbf{j}$.
The unit vector is

$$\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{\mathbf{v}}{\sqrt{3^2 + 4^2}} = \frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}\,.$$

Now

$$\begin{aligned} f_x(2, 0) &= (e^y - y\sin(xy))|_{(2,0)} = e^0 - 0 = 1 \\ f_y(2, 0) &= (xe^y - x\sin(xy))|_{(2,0)} = 2e^0 - 2 \cdot 0 = 2\,. \end{aligned}$$

Hence

$$\nabla f|_{(2,0)} = f_x(2, 0)\mathbf{i} + f_y(2, 0)\mathbf{j} = \mathbf{i} + 2\mathbf{j}$$

and so

$$D_{\mathbf{u}}f|_{(2,0)} = \nabla f|_{(2,0)} \cdot \mathbf{u} = (\mathbf{i} + 2\mathbf{j}) \cdot \left(\frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}\right) = \frac{3}{5} - \frac{8}{5} = -1\,.$$

Note that

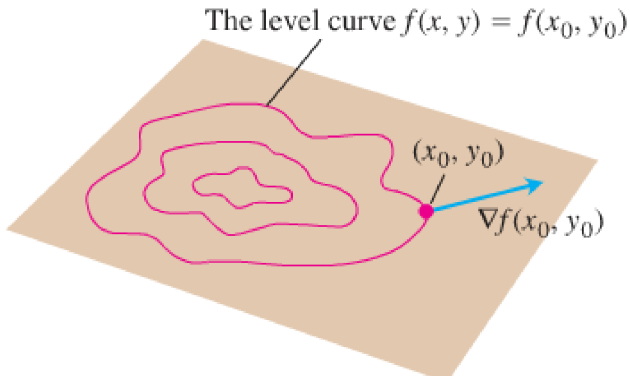
$$\boxed{D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u} = |\nabla f| \cos\theta}$$

where θ is the angle between the vectors ∇f and $\mathbf{u}$. This implies the following:

1. f increases most rapidly when $\cos\theta = 1$ (i.e. $\mathbf{u}$ is parallel to ∇f)
2. f decreases most rapidly when $\cos\theta = -1$ (i.e. $\mathbf{u}$ is in opposite direction to ∇f)
3. f has zero change when $\cos\theta = 0$ (i.e. $\mathbf{u}$ is orthogonal to ∇f).

Point 3 implies (why?):

> At every point (x_0, y_0) in the domain of a differentiable function $f(x, y)$ the gradient of f is normal to the level curve through (x_0, y_0).



Tangent lines to level curves are always normal to the gradient. If (x, y) is a point on the tangent line through the point $P(x_0, y_0)$ then

$$\mathbf{T} = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j}\,,$$

is a vector parallel to it. The *equation of the tangent* is then

$$\nabla f \cdot \mathbf{T} = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0\,.$$

Tangent Planes and Differentials

> **DEFINITIONS** Tangent Plane, Normal Line
>
> The **tangent plane** at the point $P_0(x_0, y_0, z_0)$ on the level surface $f(x, y, z) = c$ of a differentiable function f is the plane through P_0 normal to $\nabla f|_{P_0}$.
>
> The **normal line** of the surface at P_0 is the line through P_0 parallel to $\nabla f|_{P_0}$.

It follows[1] that the equation of the tangent plane is

$$\nabla f|_{P_0} \cdot \vec{P_0 P} = f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0) + f_z(P_0)(z - z_0) = 0$$

and the equation of the normal line is

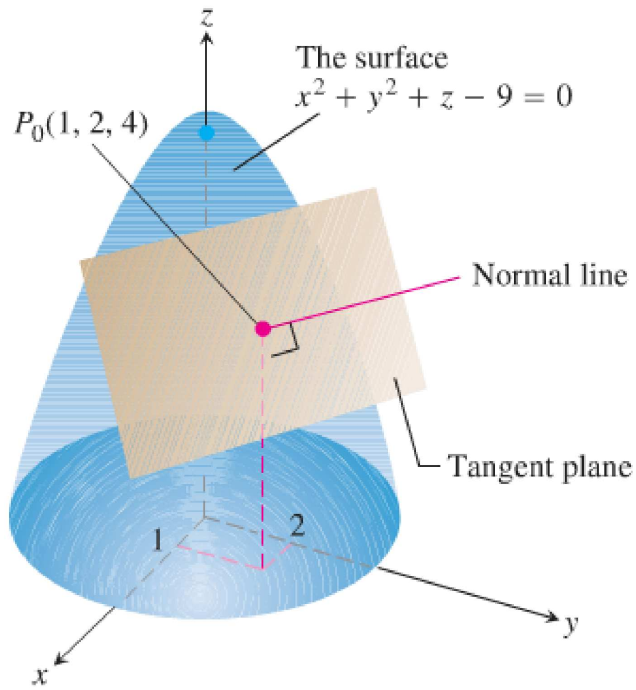
$$x = x_0 + f_x(P_0)t\,, \qquad y = y_0 + f_y(P_0)t\,, \qquad z = z_0 + f_z(P_0)t\,.$$

Example:
Find the tangent plane and normal line of the surface

$$f(x, y, z) = x^2 + y^2 + z - 9 = 0$$

(a circular paraboloid) at the point $P_0(1, 2, 4)$

[1]See Section 12.5 in Thomas' Calculus for details if you are in trouble with this.



$$\nabla f|_{P_0} = (2x\,\mathbf{i} + 2y\,\mathbf{j} + \mathbf{k})_{(1,2,4)} = 2\,\mathbf{i} + 4\,\mathbf{j} + \mathbf{k}$$

where at the point P_0 we have $f_x(P_0) = 2$, $f_y(P_0) = 4$ and $f_z(P_0) = 1$. Therefore the equation of the tangent plane is

$$2(x-1) + 4(y-2) + (z-4) = 0$$

which simplifies to

$$2x + 4y + z = 14\,.$$

The normal line to the surface at P_0 is

$$x = 1 + 2t\,, \qquad y = 2 + 4t\,, \qquad z = 4 + t\,.$$

We remark that the gradient has the following algebraic properties:

$$\begin{aligned} \nabla(kf) &= k\,\nabla f \qquad \text{for any number } k \\ \nabla(f \pm g) &= \nabla f \pm \nabla g \\ \nabla(f - g) &= \nabla f - \nabla g \\ \nabla(fg) &= f\,\nabla g + g\,\nabla f \\ \nabla\left(\frac{f}{g}\right) &= \frac{g\,\nabla f - f\,\nabla g}{g^2} \end{aligned}$$

(the proof is straightforward and is left as an exercise)
Before we linearise a function of two variables, recall that a function $z = f(x, y)$ is *differentiable* at (x_0, y_0) if

$$\Delta z = f(x,y) - f(x_0,y_0) = f_x(x_0,y_0)\Delta x + f_y(x_0,y_0)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y$$

with $\epsilon_1, \epsilon_2 \to 0$ $(\Delta x, \Delta y \to 0)$. Solve for $f(x, y)$ and approximate:

DEFINITIONS Linearization, Standard Linear Approximation

The **linearization** of a function $f(x, y)$ at a point (x_0, y_0) where f is differentiable is the function

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \qquad (5)$$

The approximation

$$f(x, y) \approx L(x, y)$$

is the **standard linear approximation** of f at (x_0, y_0).

Example:
Find the linearisation of

$$f(x, y) = x^2 - xy + \frac{1}{2}y^2 + 3$$

at the point $(3, 2)$.
We first evaluate f, f_x and f_y at the point $(x_0, y_0) = (3, 2)$:

$$\begin{aligned}
f(3,2) &= \left(x^2 - xy + \frac{1}{2}y^2 + 3\right)_{(3,2)} = 8 \\
f_x(3,2) &= \frac{\partial}{\partial x}\left(x^2 - xy + \frac{1}{2}y^2 + 3\right)_{(3,2)} = (2x - y)_{(3,2)} = 4 \\
f_y(3,2) &= \frac{\partial}{\partial y}\left(x^2 - xy + \frac{1}{2}y^2 + 3\right)_{(3,2)} = (-x + y)_{(3,2)} = -1
\end{aligned}$$

giving

$$\begin{aligned}
L(x, y) &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\
&= 8 + (4)(x - 3) + (-1)(y - 2) = 4x - y - 2\,.
\end{aligned}$$

Hence the linearisation of f at $(3, 2)$ is $L(x, y) = 4x - y - 2$.
Recall that for $y = f(x)$ we have defined the *differential* $dy = f'(x)dx$.

DEFINITION Total Differential

If we move from (x_0, y_0) to a point $(x_0 + dx, y_0 + dy)$ nearby, the resulting change

$$df = f_x(x_0, y_0)\, dx + f_y(x_0, y_0)\, dy$$

in the linearization of f is called the **total differential of f**.

Example:
The volume $V = \pi r^2 h$ of a cylinder is to be calculated from measured values of r (the radius) and h (the height). Suppose that r is measured with an error of no more than 2%

and h with an error of no more than 0.5%. Estimate the resulting possible percentage error in the calculation of V.
First note that

$$\left|\frac{dr}{r}100\right| \leq 2\,, \quad \left|\frac{dh}{h}100\right| \leq 0.5\,.$$

Then

$$dV = V_r\,dr + V_h\,dh = 2\pi rh\,dr + \pi r^2\,dh$$

and so

$$\frac{dV}{V} = \frac{2\pi rh\,dr + \pi r^2\,dh}{\pi r^2 h} = \frac{2\,dr}{r} + \frac{dh}{h}\,.$$

Hence

$$\left|\frac{dV}{V}\right| = \left|2\frac{dr}{r} + \frac{dh}{h}\right| \leq \left|2\frac{dr}{r}\right| + \left|\frac{dh}{h}\right| \leq 2(0.02) + 0.005 = 0.045\,.$$

Therefore the error is no more than 4.5%.

Extreme Values and Saddle Points

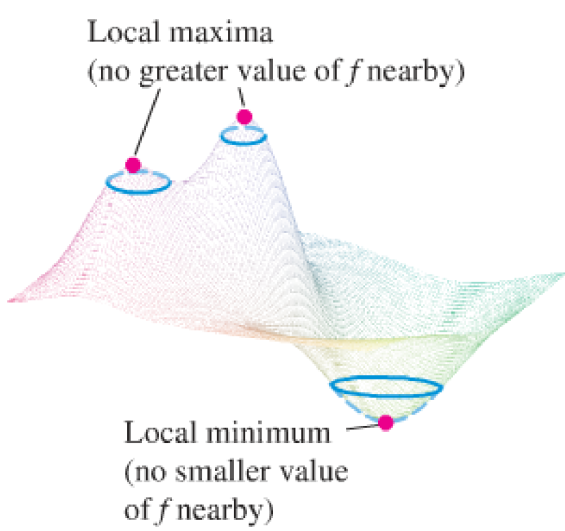
When we investigated extreme values for functions of one variable we looked for points where the graph had a horizontal tangent line. For functions of two variables we look for points where the *surface* defined by $z = f(x, y)$ has a *horizontal tangent plane*. This leads to the following definition:

DEFINITIONS **Local Maximum, Local Minimum**

Let $f(x, y)$ be defined on a region R containing the point (a, b). Then

1. $f(a, b)$ is a **local maximum** value of f if $f(a, b) \geq f(x, y)$ for all domain points (x, y) in an open disk centered at (a, b).
2. $f(a, b)$ is a **local minimum** value of f if $f(a, b) \leq f(x, y)$ for all domain points (x, y) in an open disk centered at (a, b).

Local maxima correspond to "mountain peaks" on the surface $z = f(x, y)$ and local minima correspond to "valley bottoms":



Not too hard to show:

> **THEOREM 10—First Derivative Test for Local Extreme Values** If $f(x, y)$ has a local maximum or minimum value at an interior point (a, b) of its domain and if the first partial derivatives exist there, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

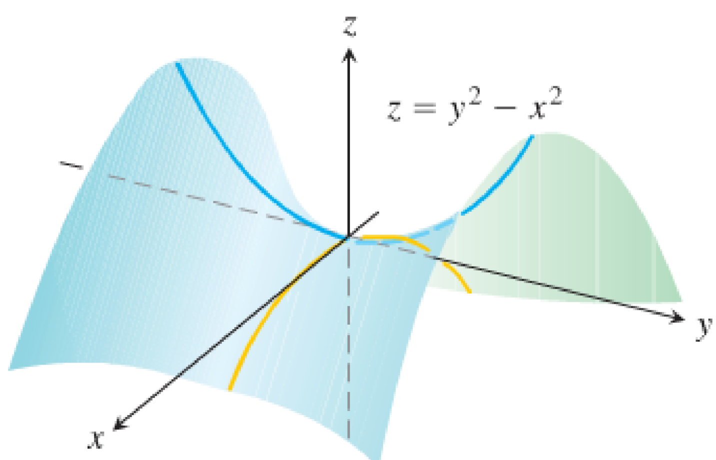
Define an important object:

> **DEFINITION Critical Point**
> An interior point of the domain of a function $f(x, y)$ where both f_x and f_y are zero or where one or both of f_x and f_y do not exist is a **critical point** of f.

Therefore local maxima and minima are critical points (why?) but critical points can also include **saddle points**:

> **DEFINITION Saddle Point**
> A differentiable function $f(x, y)$ has a **saddle point** at a critical point (a, b) if in every open disk centered at (a, b) there are domain points (x, y) where $f(x, y) > f(a, b)$ and domain points (x, y) where $f(x, y) < f(a, b)$. The corresponding point $(a, b, f(a, b))$ on the surface $z = f(x, y)$ is called a saddle point of the surface (Figure 14.40).

An example of a saddle point is the origin in the following surface:



Therefore, finding critical points of a function is not sufficient to identify the type of critical point (local maximum, local minimum or saddle point). To do this we need to make use of second partial derivatives.

THEOREM 11—Second Derivative Test for Local Extreme Values Suppose that $f(x, y)$ and its first and second partial derivatives are continuous throughout a disk centered at (a, b) and that $f_x(a, b) = f_y(a, b) = 0$. Then

i) f has a **local maximum** at (a, b) if $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).
ii) f has a **local minimum** at (a, b) if $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).
iii) f has a **saddle point** at (a, b) if $f_{xx}f_{yy} - f_{xy}^2 < 0$ at (a, b).
iv) **the test is inconclusive** at (a, b) if $f_{xx}f_{yy} - f_{xy}^2 = 0$ at (a, b). In this case, we must find some other way to determine the behavior of f at (a, b).

The quantity $f_{xx}f_{yy} - f_{xy}^2$ is called the **discriminant** or **Hessian** of the function f. In case you know already what a determinant is (otherwise you will learn this soon in Geometry 1), note that

$$f_{xx}f_{yy} - f_{xy}^2 = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} .$$

Example:

Find the local extreme values of $f(x, y) = xy - x^2 - y^2 - 2x - 2y + 4$ and determine the nature of each.

$f(x, y)$ is defined and differentiable for all points in its domain. Hence, at extreme values f_x and f_y are simultaneously zero. This gives the two equations

$$f_x = y - 2x - 2 = 0\,; \qquad f_y = x - 2y - 2 = 0\,.$$

The solution of these equations is $x = y = -2$. Hence $(-2, -2)$ is the only point where f may take an extreme value. Now take the second derivatives:

$$f_{xx} = -2 < 0\,, \qquad f_{yy} = -2\,, \qquad f_{xy} = 1\,.$$

At the point $(-2, -2)$,

$$f_{xx}f_{yy} - f_{xy}^2 = (-2)(-2) - 1^2 = 3 > 0\,.$$

So $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$. Therefore f has a local maximum at $(-2, -2)$. The value of f at this point is $f(-2, -2) = 8$.

Summary of Max-Min Tests

The extreme values of $f(x, y)$ can occur only at

i. **boundary points** of the domain of f

ii. **critical points** (interior points where $f_x = f_y = 0$ or points where f_x or f_y fail to exist).

If the first- and second-order partial derivatives of f are continuous throughout a disk centered at a point (a, b) and $f_x(a, b) = f_y(a, b) = 0$, the nature of $f(a, b)$ can be tested with the **Second Derivative Test**:

i. $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b) $\Rightarrow$ **local maximum**

ii. $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b) $\Rightarrow$ **local minimum**

iii. $f_{xx}f_{yy} - f_{xy}^2 < 0$ at (a, b) $\Rightarrow$ **saddle point**

iv. $f_{xx}f_{yy} - f_{xy}^2 = 0$ at (a, b) $\Rightarrow$ **test is inconclusive**.