Math 241
Calculus III

Thomas Honold

ZJU-UIUC Institute

Fall Semester 2020

Outline

Today's Lecture: Mean Value Theorem, Implicit Differentiation

The Mean Value Theorem

First recall the Mean Value Theorem of single-variable calculus:

If $f\colon [a,b] \to \mathbb{R}$ is continuous in $[a,b]$ and differentiable in (a,b) then there exists $\xi \in (a,b)$ such that $f(b) - f(a) = f'(\xi)(b-a)$.

Theorem (Mean Value Theorem)

Suppose $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, is differentiable and $\mathbf{a}, \mathbf{b} \in D$ are such that the line segment $[\mathbf{a}, \mathbf{b}] = \{(1-t)\mathbf{a} + t\,\mathbf{b}; 0 \leq t \leq 1\}$ is contained in D. Then there exists a point $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$ such that

$$f(\mathbf{b}) - f(\mathbf{a}) = \mathrm{d}f(\mathbf{x})(\mathbf{b} - \mathbf{a}) = \nabla f(\mathbf{x}) \cdot (\mathbf{b} - \mathbf{a}).$$

Proof.

Consider the one-variable function $\phi\colon [0,1] \to \mathbb{R}$, $t \mapsto f\big((1-t)\mathbf{a} + t\mathbf{b}\big)$, which satisfies $\phi(0) = f(\mathbf{a})$, $\phi(1) = f(\mathbf{b})$. The chain rule gives

$$\phi'(t) = \mathrm{d}f\big(\mathbf{a} + t(\mathbf{b} - \mathbf{a})\big)(\mathbf{b} - \mathbf{a}) \quad \text{for } t \in [0,1].$$

By the Mean Value Theorem of Calculus I, there exists $\xi \in (0,1)$ such $\phi(1) - \phi(0) = \phi'(\xi)$.
$\Longrightarrow$ $\mathbf{x} = \mathbf{a} + \xi(\mathbf{b} - \mathbf{a})$ has the required property. $\square$

Integral Version

As in Calculus I, there exists an integral representation of the difference $f(\mathbf{b}) - f(\mathbf{a})$. It is obtained at follows:

$$\begin{aligned} f(\mathbf{b}) - f(\mathbf{a}) &= \int_0^1 \phi'(t)\,\mathrm{d}t = \int_0^1 \mathrm{d}f\big(\mathbf{a} + t(\mathbf{b} - \mathbf{a})\big)(\mathbf{b} - \mathbf{a})\,\mathrm{d}t \\ &= \left(\int_0^1 \nabla f\big(\mathbf{a} + t(\mathbf{b} - \mathbf{a})\big)\,\mathrm{d}t\right) \cdot (\mathbf{b} - \mathbf{a}) \end{aligned}$$

In contrast with the "Lagrange (mid-point) version" this generalizes immediately to the case $m > 1$.

Mean Value Theorem (Integral Version)

Suppose $f \colon D \to \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, is differentiable and $\mathbf{a}, \mathbf{b} \in D$ are such that the line segment $[\mathbf{a}, \mathbf{b}] = \big\{(1 - t)\mathbf{a} + t\,\mathbf{b}; 0 \leq t \leq 1\big\}$ is contained in D. Then

$$f(\mathbf{b}) - f(\mathbf{a}) = \left(\int_0^1 \mathbf{J}_f\big(\mathbf{a} + t(\mathbf{b} - \mathbf{a})\big)\,\mathrm{d}t\right)(\mathbf{b} - \mathbf{a}).$$

Another Theorem of Calculus I Generalized

Recall that a function $f\colon [a,b]$ satisfying $f'(x)=0$ for $x\in[a,b]$ must be constant.

Definition

A subset $D\subseteq\mathbb{R}^n$ is said to be *path-connected* if for any two points $\mathbf{a},\mathbf{b}\in D$ there exists a continuous curve $g\colon [0,1]\to D$ satisfying $g(0)=\mathbf{a}$, $g(1)=\mathbf{b}$. In other words, one can continuously walk from $\mathbf{a}$ to $\mathbf{b}$ without ever leaving D.

Examples

Convex sets (i.e, sets $D\subseteq\mathbb{R}^n$ which contain with any two points $\mathbf{a},\mathbf{b}$ also the straight-line segment $[\mathbf{a},\mathbf{b}]$) are path-connected. A further example is $\mathbb{R}^n\setminus\{\mathbf{0}\}$ for $n\geq 2$, which shows that path-connected sets may contain “holes”.

Theorem

Suppose $D\subseteq\mathbb{R}^n$ is path-connected and $f\colon D\to\mathbb{R}^m$ is differentiable with $\mathrm{d}f(\mathbf{x})=0$ for every $\mathbf{x}\in D$. Then f is constant, i.e., there exists $\mathbf{c}\in\mathbb{R}^m$ such that $f(\mathbf{x})=\mathbf{c}$ for all $\mathbf{x}\in D$.

Proof.

Writing $f = (f_1, \dots, f_m)$, the condition $f(\mathbf{x}) = \mathbf{c}$ is equivalent to $f_i(\mathbf{x}) = c_i$ with $c_i \in \mathbb{R}$. Hence it suffices to consider the case $m = 1$.

Fix $\mathbf{a} \in D$ and let $c = f(\mathbf{a})$. We must show $f(\mathbf{b}) = c$ for all $\mathbf{b} \in D$.

By assumption there exists a continuous curve $g\colon [0,1] \to D$ with $g(0) = \mathbf{a}$, $g(1) = \mathbf{b}$. Since every point $g(t)$ is an inner point of D (because f is assumed to be differentiable at $g(t)$), we can smooth g, if necessary, and assume that g' exists. Then we can apply the theorem of Calculus I to the composition $\phi\colon [0,1] \to \mathbb{R}$, $t \mapsto f(g(t))$, which has derivative $\phi'(t) = \mathrm{d}f\big(g(t)\big)\big(g'(t)\big) = 0$, and conclude that $f(\mathbf{b}) = \phi(1) = \phi(0) = f(\mathbf{a})$. $\square$

And Yet another One

Theorem (Intermediate Value Theorem)

Suppose $D \subseteq \mathbb{R}^n$ is open and path-connected and $f\colon D \to \mathbb{R}$ is continuous. Then the range $f(D)$ is an interval.

This theorem holds, more generally, for continuous real-valued functions whose domain $D \subseteq \mathbb{R}^n$ is *connected* (but not necessarily open) and says that f must attain every value between any two given values $f(\mathbf{a})$ and $f(\mathbf{b})$.

Proof.

Connect $\mathbf{a}$ and $\mathbf{b}$ by a continuous curve $g\colon [0,1] \to D$ and apply the intermediate value theorem for one-variable functions to $[0,1] \to \mathbb{R}$, $t \mapsto f\bigl(g(t)\bigr)$, which is continuous since it is a composition of continuous functions. $\square$

Error Propagation

Especially when using floating-point computations

In real-world applications we are often faced with evaluating a real-valued function $y = f(x_1, \dots, x_n)$ for some input data $x_1, \dots, x_n$ which is not accurately known.

Question

How do small errors in the input affect the output y ?

An answer can be given using differentials. If the true input x_j is replaced by $x_j + \Delta x_j$, the corresponding change in the output y is

$$\begin{aligned}\Delta y &= f(x_1 + \Delta x_1, \dots, x_n + \Delta x_n) - f(x_1, \dots, x_n) \\ &\approx \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_1, \dots, x_n)\Delta x_j \quad \text{if the } \Delta x_j \text{ are small.}\end{aligned}$$

The Mean Value Theorem provides a rigorous answer:

$$\Delta y = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x} + \tau \Delta \mathbf{x})\Delta x_j \quad \text{for some } \tau \in [0, 1],$$

were $\mathbf{x} = (x_1, \dots, x_n)$ and $\Delta \mathbf{x} = (\Delta x_1, \dots, \Delta x_n)$.

Usually we do not know the exact value of Δx_j but only a bound $|\Delta x_j| \leq \delta_j$. In this case the Mean Value Theorem yields the upper bound

$$|\Delta y| \leq \sum_{j=1}^{n} M_j \delta_j \quad \text{with} \quad M_j = \max_{\mathbf{x}' \in U} \left| \frac{\partial f}{\partial x_j}(\mathbf{x}') \right|,$$

where $U = \{\mathbf{x}' \in \mathbb{R}^n; \left| x_j' - x_j \right| \leq \delta_j \text{ for } 1 \leq j \leq n\}$ is the "uncertainty region" of the input.

Example (taken from [Ste16], p. 934)

The dimensions of a rectangular box are measured as 75 cm, 60 cm, and 40 cm. Each measurement is correct within 0.2 cm. Estimate the largest possible error when the volume of the box is calculated from these measurements.

Solution: We use cm as the unit of measurement.

The differential of the volume $V = V(x, y, z) = xyz$ is

$$\mathrm{d}V = \mathrm{d}V(x, y, z) = yz\,\mathrm{d}x + xz\,\mathrm{d}y + xy\,\mathrm{d}z,$$

giving the error in the volume computation as

$$\Delta V \approx \mathrm{d}V(x, y, z)(\Delta x, \Delta y, \Delta z) = yz\,\Delta x + xz\,\Delta y + xy\,\Delta x.$$

Example (cont'd)

Clearly the partial derivatives take their maximum in the uncertainty region at $(75.2, 60.2, 40.2)$.

$$\begin{aligned}\implies |\Delta V| &\leq (60.2)(40.2)(0.2) + (75.2)(40.2)(0.2) + (60.2)(75.2)(0.2) \\ &= 1994.024\end{aligned}$$

The maximum possible error in the volume computation is therefore $1994.024\,\mathrm{cm}^3$.

Note: We have used the fact that the uncertainty region can also be expressed as $U = \{\mathbf{x}' \in \mathbb{R}^n;\, \left|x_j' - x_j - \Delta x_j\right| \leq \delta_j \text{ for } 1 \leq j \leq n\}$.

Relative Errors

For practical purposes it is more useful to determine the propagation of relative errors $\Delta x_j / x_j$, because floating-point computations cause additional errors, whose relative size can be controlled. The preceding considerations give

$$\frac{\Delta y}{y} \approx \sum_{j=1}^{n} \frac{x_j}{y} \frac{\partial f}{\partial x_j}(x_1, \dots, x_n) \frac{\Delta x_j}{x_j} \quad \text{for small } \Delta_j,$$

and a corresponding exact version.

Example (cont'd)

The relative error of the volume computation is

$$\begin{aligned}\frac{\Delta V}{V} &\approx \frac{\Delta x}{x}+\frac{\Delta y}{y}+\frac{\Delta z}{z}\\ &= \frac{1}{5}\left(\frac{1}{75}+\frac{1}{60}+\frac{1}{40}\right)=0.011,\end{aligned}$$

or about 1 %.

Remark

The preceding example is a toy example and it is easy to determine the maximum/minimum absolute error directly:

$$\begin{aligned}\Delta V &\leq (75.2)(60.2)(40.2)-65\cdot 60\cdot 40\approx 1987,\\ \Delta V &\geq (74.8)(59.8)(39.8)-65\cdot 60\cdot 40\approx -1973.\end{aligned}$$

Perhaps it is instructive to see how the bound resulting from the Mean Value Theorem overestimates the error (with $h=0.2$):

$$\begin{aligned}|\Delta V| &\leq (y+h)(z+h)h+(x+h)(z+h)h+(x+h)(y+h)h\\ &= (yz+xz+xy)h+(x+y+z)2h^2+3h^3,\end{aligned}$$

Remark (cont'd)

whereas direct expansion gives

$$\begin{aligned}|\Delta V| &\leq (x+h)(y+h)(z+h) - xyz \\ &= (yz+xz+xy)h + (x+y+z)h^2 + h^3.\end{aligned}$$

In most cases, however, direct expansion of the absolute error is impossible (too complicated), and the bound resulting from the Mean Value Theorem is a viable alternative.

Floating-point numbers

A t-digit *floating-point number (fpn)* is a real number of the form $\pm m \times b^e$, where $b \geq 2$ is a fixed integer (the *base*), e is an integer from some finite range $\mathrm{e}_{\min} \leq e \leq \mathrm{e}_{\max}$ (*exponent*) and $m > 0$ (*mantissa*) is a number that has a base-b representation using t digits.

Notes

- The representation is usually normalized to $m \in [1, b)$, so that the base-b representation of m has the form $m_1.m_2m_3\dots m_t = \sum_{i=1}^{t} m_i b^{1-i}$ with $1 \leq m_1 \leq b-1$ and $0 \leq m_i \leq b-1$ for $i \geq 2$.
- The number b^{1-t}, which represents the distance from 1 to the smallest fpn > 1, is called *machine epsilon* and denoted by ε_{M}.

Example

Double-precision floating-point approximations to π and π^{100} in SageMath :

$$\pi \approx \texttt{3.14159265358979}, \quad \pi^{100} \approx \texttt{5.18784831431959e49}$$

They are written as 15-digit decimal fpn on the screen, but the internal representation uses binary 53-digit fpn according to the IEEE double-precision standard.

Fact

The standard arithmetic functions (including square roots, exp, log, sin, cos, etc.) can be implemented in such a way that the result $\hat{y}$ of a computation equals the exact result y rounded to the nearest floating-point number (except when overflow/underfow occurs).

This gives for the relative error $\epsilon = \frac{\hat{y}-y}{y}$ the bound $\epsilon \leq \varepsilon_{\mathrm{M}}$, or

$$\hat{y} = y(1+\epsilon) \quad \text{with} \quad |\epsilon| \leq \varepsilon_{\mathrm{M}}.$$

Observation

For iterated computations the error in the output of one computation acts as error in the input data of the next computation, and these errors propagate to errors in the final result.

Example

When computing the sum of 3 numbers x_1, x_2, x_3 in floating-point arithmetic, we have

$$\begin{aligned}
\widehat{x_1 + x_2} &= (x_1 + x_2)(1+\epsilon_1), && (|\epsilon_1| \leq \varepsilon_{\mathrm{M}})\\
x_1 \,\widehat{+\, x_2 \,+}\, x_3 &= \big((x_1+x_2)(1+\epsilon_1) + x_3\big)(1+\epsilon_2) && (|\epsilon_2| \leq \varepsilon_{\mathrm{M}})\\
&= x_1(1+\epsilon_1)(1+\epsilon_2) + x_2(1+\epsilon_1)(1+\epsilon_2) + x_3(1+\epsilon_2)
\end{aligned}$$

Remarks

- The modern point-of-view emphasizes so-called *backwards analysis*: The approximate result of a machine computation is exact for (slightly) changed input data. For example, the previous computation returned the exact result of adding the 3 numbers $x_1(1+\epsilon_1)(1+\epsilon_2)$, $x_2(1+\epsilon_1)(1+\epsilon_2)$ and $x_3(1+\epsilon_2)$. Since these differ from x_1, x_2, x_3 only by a small relative error ($\leq 2\varepsilon_{\mathrm{M}}$), the operation is *well-conditioned*.
- Of the 4 basic arithmetic operations, only subtraction $s(x,y) = x - y$ is ill-conditioned (and only if the numbers involved have approximately the same size).
  In order to understand this, we estimate the propagation of the relative error: Since $\mathrm{d}s = \mathrm{d}x - \mathrm{d}y$, we have
  $$\Delta s \approx \Delta x - \Delta y,$$
  $$\frac{\Delta s}{s} \approx \frac{x}{x-y}\frac{\Delta x}{x} - \frac{y}{x-y}\frac{\Delta y}{y}.$$
  It is an instructive exercise to compute $1 - 0.999999$ in 6-digit decimal floating-point arithmetic and compare it with the exact result.

Implicit Differentiation

For a differentiable function $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^2$, the gradient $\nabla f(\mathbf{x})$ is perpendicular to the contour line at every point $\mathbf{x} \in D$, as we have seen. Virtually the same proof shows that in the general case $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ the gradient $\nabla f(\mathbf{x})$ (provided it is nonzero) serves as a normal vector to the contour hypersurface of f at $\mathbf{x}$. Moreover we can compute the partial derivatives (and hence the differential) of any function g implicitly defined by the contour hypersurface. For this we consider a 3-dimensional example, which is a slight variant of one in our textbook [Ste16].

Example

Find the partial derivatives of the function $z = g(x, y)$ defined implicitly by the 2-dimensional surface S with equation

$$x^3 + y^3 + z^3 + 3xyz = 1 \quad \text{in } \mathbb{R}^3.$$

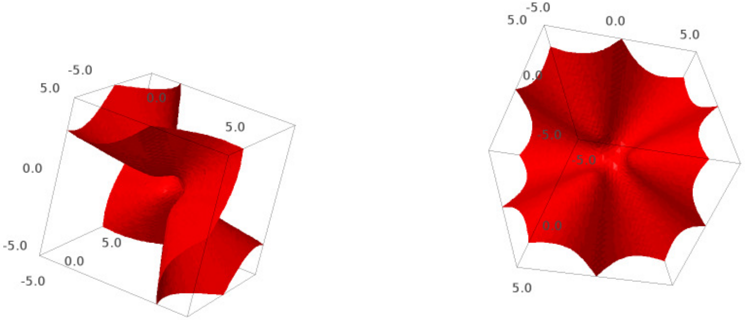

Figure: The surface S from different perspectives

The second picture reveals that S is invariant under a certain rotation of $\mathbb{R}^3$ by an angle of 120°.

Example (cont'd)

Let $F(x,y,z) = x^3 + y^3 + z^3 + 3xyz - 1$. Then the required function g must satisfy $F\bigl(x,y,g(x,y)\bigr) = 0$.

Applying the Chain Rule to F and $G(x,y) = \bigl(x,y,g(x,y)\bigr)^{\mathsf{T}}$ (recall that vector-valued functions must be written as column vectors when computing their Jacobi matrices!) gives, writing $z = g(x,y)$ and later suppressing arguments,

$$
\begin{aligned}
& 0 = (F \circ G)(x,y), && \text{(for all (x,y) in the domain of g)} \\
\implies & 0 = \mathrm{d}F\bigl(x,y,g(x,y)\bigr) \circ \mathrm{d}G(x,y), && \text{(zero linear map)} \\
\implies & \mathbf{0} = \mathbf{J}_F\bigl(x,y,g(x,y)\bigr)\mathbf{J}_G(x,y) && \text{(all-zero matrix)} \\
& = \begin{pmatrix} F_x & F_y & F_z \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ g_x & g_y \end{pmatrix} = \begin{pmatrix} F_x + F_z g_x & F_y + F_z g_y \end{pmatrix}.
\end{aligned}
$$

$$
\begin{aligned}
\implies g_x &= -\frac{F_x}{F_z} = -\frac{3x^2 + 3yz}{3z^2 + 3xy} = -\frac{x^2 + yz}{z^2 + xy}, \\
g_y &= -\frac{F_y}{F_z} = -\frac{3y^2 + 3xz}{3z^2 + 3xy} = -\frac{y^2 + xz}{z^2 + xy}.
\end{aligned}
$$

Example (cont'd)

How to find the domain of g?

A qualitative answer is provided by the

Implicit Function Theorem (special case):
Suppose $P_0 = (x_0, y_0, z_0)$ satisfies $F(x_0, y_0, z_0) = 0$ and $F_z(x_0, y_0, z_0) \neq 0$, i.e., P_0 is on the surface and the normal vector to the surface at P_0 is not contained in the x, y-plane. Then there exists a neighborhood of P_0 of the form $U' \times U''$ with

$$\begin{aligned} U' &= \left\{(x, y) \in \mathbb{R}^2; |x - x_0| < \delta' \wedge |y - y_0| < \delta'\right\}, \\ U'' &= \left\{z \in \mathbb{R}; |z - z_0| < \delta''\right\}, \qquad \delta', \delta'' > 0, \end{aligned}$$

and a C^1-function $g \colon U' \to U''$ such that the part of the surface contained in $U' \times U''$ is precisely the graph of g, i.e.,

$$F(x, y, z) = 0 \wedge (x, y, z) \in U' \times U'' \iff z = g(x, y)$$

This theorem, which mutatis mutandis applies to any C^1-function $F \colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, and any of its n variables (in place of z), is a mere existence result and doesn't provide any explicit description of U', U'', and g.

Example (cont'd)

Now we determine an appropriate neighborhood $U' \times U''$ for the point $P_0 = (0,0,1)$, which satisfies the condition $F_z(0,0,1) = 3 \neq 0$.

For fixed x, y consider the function

$$h(z) = h_{x,y}(z) = F(x,y,z) = x^3 + y^3 + z^3 + 3xyz - 1.$$

We have

$$h'(z) = F_z(x,y,z) = 3(z^2 + xy) > 0 \quad \text{if } |x| < \tfrac{1}{2}, |y| < \tfrac{1}{2}, |z-1| < \tfrac{1}{2}.$$

$\Longrightarrow h = h_{x,y}$ is strictly increasing on $\left(\frac{1}{2}, \frac{3}{2}\right)$, provided we restrict (x,y) to $U' = \left\{(x,y) \in \mathbb{R}^2; |x| < \frac{1}{2}, |y| < \frac{1}{2}\right\}$. Moreover, for those x, y, z we have

$$h\left(\tfrac{1}{2}\right) = x^3 + y^3 + \tfrac{3}{2}xy - \tfrac{7}{8} \leq \tfrac{1}{8} + \tfrac{1}{8} + \tfrac{3}{2} \cdot \tfrac{1}{4} - \tfrac{7}{8} < 0,$$
$$h\left(\tfrac{3}{2}\right) = x^3 + y^3 + \tfrac{9}{2}xy + \tfrac{19}{8} \geq -\tfrac{1}{8} - \tfrac{1}{8} - \tfrac{9}{2} \cdot \tfrac{1}{4} + \tfrac{19}{8} > 0,$$

showing that the unique solution of $h(z) = 0$ falls into the interval $\left(\frac{1}{2}, \frac{3}{2}\right)$. $\Longrightarrow$ We can take U' as above and $U'' = \left(\frac{1}{2}, \frac{3}{2}\right)$.

Example (cont'd)

Note that a similar argument will work for any C^1-function $F\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$.

It remains to show that g is continuously differentiable. This can be done as follows. Applying the Mean Value Theorem to F, we have for any $(x, y) \in U'$ and s, t sufficiently small an identity

$$\begin{aligned} 0 &= F\bigl(x+s, y+t, g(x+s, y+t)\bigr) - F\bigl(x, y, g(x, y)\bigr) \\ &= F_x(\xi,\eta,\zeta)s + F_y(\xi,\eta,\zeta)t + F_z(\xi,\eta,\zeta)\bigl(g(x+s,y+t) - g(x,y)\bigr), \end{aligned}$$

where (ξ,η,ζ) is a point between $\bigl(x, y, g(x, y)\bigr)$ and $\bigl(x+s, y+t, g(x+s, y+t)\bigr)$. Since F_x, F_y, and $1/F_z$ are continuous on $U' \times U''$, they are bounded, and hence letting $(s, t) \to (0, 0)$ in the above identity gives $g(x+s, y+t) \to g(x, y)$, i.e., the continuity of g. Moreover, setting $t = 0$ we get

$$\frac{g(x+s, y) - g(x, y)}{s} = -\frac{F_x(\xi,\eta,\zeta)}{F_z(\xi,\eta,\zeta)} \to -\frac{F_x\bigl(x, y, g(x, y)\bigr)}{F_z\bigl(x, y, g(x, y)\bigr)} \quad \text{for } s \to 0.$$

$\Longrightarrow$ $\frac{\partial g}{\partial x}$ (and similarly $\frac{\partial g}{\partial y}$) exists and is continuous, because it is composed of continuous functions. $\Longrightarrow$ g is a C^1-function.

Example (cont'd)

Why is S "smooth"?

The gradient ∇F doesn't vanish at any point of S, since the only solution of

$$\nabla F(x,y,z) = 0 \iff x^2 + yz = y^2 + xz = z^2 + xy = 0$$

is $(0,0,0)$, but $(0,0,0) \notin S$.
$\implies$ At every point of S at least one of the representations $z = g(x,y)$, $y = h(x,z)$ or $x = k(y,z)$ exists.
$\implies$ S has a tangent plane at every point.

How to obtain the tangent plane?

The tangent plane to S in $P_0 = (x_0, y_0, z_0)$ has normal vector $\nabla F(x_0, y_0, z_0)$ and hence the equation

$$F_x(x_0,y_0,z_0)(x-x_0)+F_y(x_0,y_0,z_0)(y-y_0)F_z(x_0,y_0,z_0)(z-z_0) = 0.$$

For example, $\nabla F(0,0,1) = (0,0,1)$ gives for the tangent plane at $P_0 = (0,0,1)$ the equation $0 \cdot x + 0 \cdot y + 3(z-1) = 0$ or, simplified, $z - 1 = 0$.

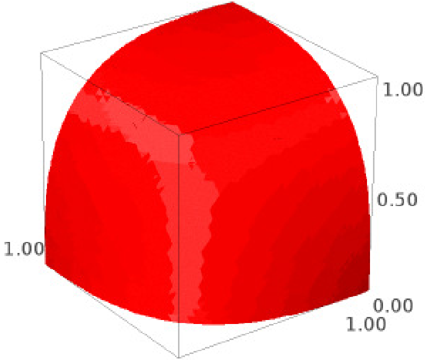

Figure: Intersection of S with the unit cube in $\mathbb{R}^3$

Higher Derivatives

If $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, is partially differentiable, then $f_{x_j} = \frac{\partial f}{\partial x_j}$ also has domain D and codomain $\mathbb{R}$. Asking whether the partial derivatives are itself (partially) differentiable, leads to the notion of k-th order partial derivatives for $k = 1, 2, 3, \dots$ (as in Calculus I).

It will be convenient to consider partial differentiation with respect to x_j as an *operator* on the set of functions $f\colon D \to \mathbb{R}$, because higher derivatives then correspond to iterating such operators.

Notation

- For 1st-order partial derivatives we use the following additional notation: $\mathrm{D}_j f = \frac{\partial}{\partial x_j} f = \frac{\partial f}{\partial x_j} = f_{x_j}$.
  Thus $\mathrm{D}_j = \frac{\partial}{\partial x_j}$ is considered as a map from functions to functions with the same (or smaller) domain.
- For 2nd-order partial derivatives we write

$$\mathrm{D}_1\mathrm{D}_1 f = \mathrm{D}_1(\mathrm{D}_1 f) = \frac{\partial^2 f}{\partial x_1^2} = (f_{x_1})_{x_1} = f_{x_1 x_1},$$

$$\mathrm{D}_2\mathrm{D}_1 f = \mathrm{D}_2(\mathrm{D}_1 f) = \frac{\partial^2 f}{\partial x_2 \partial x_1} = (f_{x_1})_{x_2} = f_{x_1 x_2}, \quad \textit{etc.}$$

Example

We compute the higher-order partial derivatives of $f(x,y) = x^3 - 3xy^2$.

$$f_x = 3x^2 - 3y^2, \qquad f_y = -6xy$$

$$f_{xx} = 6x, \qquad f_{xy} = -6y$$
$$f_{yx} = -6y, \qquad f_{yy} = -6x$$

$$f_{xxx} = 6, \qquad f_{xxy} = 0$$
$$f_{xyx} = 0, \qquad f_{xyy} = -6$$
$$f_{yxx} = 0, \qquad f_{yxy} = -6$$
$$f_{yyx} = -6, \qquad f_{yyy} = 0$$

One observes that the order in which the operators D_x and D_y are applied does not matter, viz. $f_{xy} = f_{yx}$, $f_{xxy} = f_{xyx} = f_{yxx}$, $f_{xyy} = f_{yxy} = f_{yyx}$. This is no coincidence!

Theorem (Clairaut's Theorem)

Suppose $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^2$, has continuous second-order partial derivatives. Then $f_{xy} = f_{yx}$.

An equivalent statement is the following: If f is partially differentiable in some open disk around (a, b) and the partial derivatives f_{xy}, f_{yx} are continuous on the whole disk, then $f_{xy}(a, b) = f_{yx}(a, b)$.

Corollary

Suppose $f\colon D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$, has continuous k-th order partial derivatives. Then

$$\mathrm{D}_{j_1}\mathrm{D}_{j_2}\cdots\mathrm{D}_{j_k} f = \mathrm{D}_{j_{\pi(1)}}\mathrm{D}_{j_{\pi(2)}}\cdots\mathrm{D}_{j_{\pi(k)}}$$

for any $(j_1, j_2, \ldots, j_k) \in \{1, \ldots, n\}^k$ and any permutation π of $\{1, 2, \ldots, k\}$.

Hence we can write $\mathrm{D}_{j_1}\mathrm{D}_{j_2}\cdots\mathrm{D}_{j_k} = D_1^{\alpha_1} D_2^{\alpha_2}\cdots D_n^{\alpha_n}$, where α_j denotes the number of indexes $s \in \{1, \ldots, k\}$ such that $j_s = j$.

Proof of Clairaut's Theorem.

The limit

$$L=\lim_{h\to 0}\frac{f(a+h,b+h)-f(a,b+h)-f(a+h,b)+f(a,b)}{h^2}$$

is a 2-dimensional analogon of the limit used to define derivatives in Calculus I. We evaluate this limit in two different ways, using the Mean Value Theorem of Calculus I.

1. Consider $g_1(x)=f(x,b+h)-f(x,b)$:

$$\begin{aligned}L&=\lim_{h\to 0}\frac{g_1(a+h)-g_1(a)}{h^2}=\lim_{h\to 0}\frac{g_1'(\xi_1)h}{h^2}\\&=\lim_{h\to 0}\frac{f_x(\xi_1,b+h)-f_x(\xi_1,b)}{h}=\lim_{h\to 0}f_{xy}(\xi_1,\eta_1)=f_{xy}(a,b)\end{aligned}$$

2. Consider $g_2(y)=f(a+h,y)-f(a,y)$:

$$\begin{aligned}L&=\lim_{h\to 0}\frac{g_2(b+h)-g_2(b)}{h^2}=\lim_{h\to 0}\frac{g_2'(\eta_2)}{h}\\&=\lim_{h\to 0}\frac{f_y(a+h,\eta_2)-f_y(a,\eta_2)}{h}=\lim_{h\to 0}f_{yx}(\xi_2,\eta_2)=f_{yx}(a,b)\end{aligned}$$

□

Partial Differential Equations

Partial differential equations or *PDE*'s, for short, are equations relating the partial derivatives (possibly of higher order and/or order zero) of a function. Often (but not always) they express a law of Physics, like in the following two examples.

Laplace's Equation

The 2-dimensional Laplace equation is

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

This may also be expressed as

$$\Delta u = (\mathrm{D}_1^2 + \mathrm{D}_2^2)(u) = \mathrm{D}_1(\mathrm{D}_1 u) + \mathrm{D}_2(\mathrm{D}_2 u) = 0$$

with $\Delta = \mathrm{D}_1^2 + \mathrm{D}_2^2$. The differential operator $\Delta = \mathrm{D}_1^2 + \mathrm{D}_2^2$, or $\Delta = \mathrm{D}_1^2 + \cdots + \mathrm{D}_n^2$ in general, is called *Laplace operator*. Solutions of the Laplace equation $\Delta u = 0$ are called *harmonic functions*.

Example

Show that $u(x, y) = \mathrm{e}^x \cos y$, $(x, y) \in \mathbb{R}^2$, is a solution of the 2-dimensional Laplace equation.

$$u_x = \mathrm{e}^x \cos y = u_{xx}$$

$$u_y = -\mathrm{e}^x \sin y, \qquad u_{yy} = -\mathrm{e}^x \cos(y)$$

$$\Longrightarrow u_{xx} + u_{yy} = 0$$

Note

Since partial derivatives and linear combinations (with constant coefficients!) of solutions of $\Delta u = 0$ are again solutions, all functions of the form

$$u(x, y) = A\,\mathrm{e}^x \cos y + B\,\mathrm{e}^x \sin y, \qquad A, B \in \mathbb{R}$$

are solutions of $\Delta u = 0$ (i.e., harmonic).

Example (cont'd)

Question: How can one find such solutions in the first place?

The idea is to try the „Ansatz“

$$u(x,y) = f(x)g(y) \quad \text{for some one-variable functions } f, g.$$

$$\implies u_{xx} + u_{yy} = f''(x)g(y) + f(x)g''(y) \stackrel{!}{=} 0.$$

At points (x, y) with $u(x, y) \neq 0$ this is equivalent to

$$\frac{f''(x)}{f(x)} = -\frac{g''(y)}{g(y)}.$$

Since the right-hand side is independent of x, we obtain $f''(x) = k\,f(x)$, $g''(y) = -k\,g(y)$ for some constant $k \in \mathbb{R}$ and all x, y in the domain of f resp. g, which are assumed to be intervals. Zeros of f or g make no exception.

For $k = 1$ the solutions of these differential equations are $f(x) = c_1\mathrm{e}^x + c_2\mathrm{e}^{-x}$, $g(y) = c_3\cos(y) + c_4\sin(y)$.
$\implies (x,y) \mapsto \mathrm{e}^{\pm x}\cos(y), \mathrm{e}^{\pm x}\sin(y)$ and all linear combinations thereof solve the Laplace equation. Other values of k yield further solutions.

The Preceding Example Generalized

The functions $\mathrm{e}^x \cos y$ and $\mathrm{e}^x \sin y$ arise as real and imaginary part of the complex exponential function $z \mapsto \mathrm{e}^z$:

$$\mathrm{e}^z = \mathrm{e}^{x+\mathrm{i}y} = \mathrm{e}^x \mathrm{e}^{\mathrm{i}y} = \mathrm{e}^x(\cos y + \mathrm{i}\sin y) = \mathrm{e}^x \cos y + \mathrm{i}\,\mathrm{e}^x \sin y$$

They are harmonic because $z \mapsto \mathrm{e}^z$ has a complex derivative:

$$\lim_{\substack{h\to 0\\ h\in\mathbb{C}}} \frac{\mathrm{e}^{z+h} - \mathrm{e}^z}{h} = \lim_{h\to 0} \frac{\mathrm{e}^z\mathrm{e}^h - \mathrm{e}^z}{h} = \mathrm{e}^z \lim_{h\to 0} \frac{\mathrm{e}^h - 1}{h} = \mathrm{e}^z.$$

Definition

A function $f\colon D \to \mathbb{C}$ on an open set $D \subseteq \mathbb{C}$ is said to be *holomorphic* if $f'(z) = \lim_{h\to 0} \frac{f(z+h)-f(z)}{h}$ exists for every $z \in D$.

Examples

$z \mapsto \mathrm{e}^z$; polynomials and rational functions with coefficients in $\mathbb{C}$; power series within their circle of convergence.

Notes

- The theory of holomorphic functions, called *Complex Analysis*, is fundamentally different from the theory of differentiable functions of one real variable. For example, one can show that for a holomorphic function f the derivative f' is again holomorphic, implying the existence of all derivatives of higher order!
- Since $\mathbb{C} \triangleq \mathbb{R}^2$ via $x + \mathrm{i}y \mapsto (x, y)$, we can ask about the relation between holomorphic functions and differentiable functions $\mathbb{R}^2 \to \mathbb{R}^2$. The answer is as follows:

  If $f = (u, v) = u + \mathrm{i}v$ (i.e., $u = \operatorname{Re} f$, $v = \operatorname{Im} f$) is holomorphic, we have $\mathrm{d}f(z) = f'(z)\,\mathrm{d}z$. In particular, f and hence its component functions u, v are differentiable, and

$$f_x(z) = \mathrm{d}f(z)(\mathbf{e}_1) = \mathrm{d}f(z)(1) = f'(z),$$
$$f_y(z) = \mathrm{d}f(z)(\mathbf{e}_2) = \mathrm{d}f(z)(\mathrm{i}) = \mathrm{i}f'(z).$$

  Since $f_x = (u_x, v_x) = u_x + \mathrm{i}v_x$, and similarly for f_y, this shows

$$u_y + \mathrm{i}v_y = \mathrm{i}(u_x + \mathrm{i}v_x) = \mathrm{i}u_x - v_x, \quad \text{i.e.,} \quad u_x = v_y \wedge u_y = -v_x.$$

Notes cont'd

- (item cont'd)
  $\Longrightarrow$ The Jacobi matrix $\mathbf{J}_f = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}$ has the form $\begin{pmatrix} u_x & -v_x \\ v_x & u_x \end{pmatrix}$.
  Conversely, if $f = (u, v)$ is differentiable and its partial derivatives satisfy the conditions $u_x = v_y$, $u_y = -v_x$ then

$$\begin{aligned} \mathrm{d}f(z)(h_1 + \mathrm{i}h_2) &= \big(u_x(z)h_1 + u_y(z)h_2, v_x(z)h_1 + v_y(z)h_2\big) \\ &= \big(u_x(z)h_1 - v_x(z)h_2, v_x(z)h_1 + u_x(z)h_2\big) \\ &= \big(u_x(z) + \mathrm{i}v_x(z)\big)(h_1 + \mathrm{i}h_2) \end{aligned}$$

i.e., $\mathrm{d}f(z)$ is multiplication by $u_x(z) + \mathrm{i}v_x(z)$, and

$$\begin{aligned} \frac{f(z+h) - f(z)}{h} = \frac{\mathrm{d}f(z)(h) + \mathrm{o}(h)}{h} &= u_x(z) + \mathrm{i}v_x(z) + \frac{\mathrm{o}(h)}{h} \\ &\longrightarrow u_x(z) + \mathrm{i}v_x(z) \quad \text{for } h \to 0, \end{aligned}$$

so that f is holomorphic with $f'(z) = u_x(z) + \mathrm{i}v_x(z)$.

Notes cont'd

- If $f = (u, v) = u + \mathrm{i}v$ is holomorphic then u and v are harmonic. For u, e.g., this follows easily from

$$\begin{aligned} u_{xx} &= (u_x)_x = (v_y)_x = v_{yx}, \\ u_{yy} &= (u_y)_y = (-v_x)_y = -v_{xy}, \end{aligned}$$

  and Clairaut's Theorem.

  Here we are assuming that u (or v) has continuous 2nd-order partial derivatives, which is required for the application of Clairaut's Theorem. This can in fact be shown, but it is not at all easy. (Essentially this property is equivalent to the surprising "if f is holomorphic then f' is holomorphic".)

Example

It is an instructive exercise to determine which polynomial functions $p(x,y)$ in two variables of small degree d are harmonic. Since Δ maps monomials of degree n to monomials of degree $n-2$, it suffices to consider this question for homogeneous polynomials of degree d.

$d \leq 1$ Polynomials of degree ≤ 1 are obviously harmonic.

$d = 2$ You can check that $p(x,y) = ax^2 + bxy + cy^2$ is harmonic iff $c = -a$; i.e., iff $p(x,y)$ is a linear combination of $x^2 - y^2$ and xy.

$d = 3$ We do this case in detail. For $p(x,y) = ax^3 + bx^2y + cxy^2 + dy^3$ we have

$$\begin{aligned} p_x &= 3ax^2 + 2bxy + cy^2, & p_y &= bx^2 + 2cxy + 3dy^2, \\ p_{xx} &= 6ax + 2by, & p_{yy} &= 2cx + 6dy. \end{aligned}$$

It follows that $\Delta p = (6a + 2c)x + (2b + 6d)y = 0 \iff c = -3a \wedge b = -3d$; i.e., $p(x,y) = a(x^3 - 3xy^2) + d(y^3 - 3x^2y)$.

Compare this with $z^2 = (x + y\mathrm{i})^2 = x^2 - y^2 + 2xy\,\mathrm{i}$, $z^3 = (x + y\mathrm{i})^3 = x^3 - 3xy^2 + (3x^2y - y^3)\mathrm{i}$.

Exercise

Show that a homogeneous polynomial $p(x, y)$ of degree 4 is harmonic iff $p(x, y)$ is a linear combination of $x^4 - 6x^2y^2 + y^4$ and $x^3y - y^3x$. Compare this with $\mathrm{Re}(z^4)$ and $\mathrm{Im}(z^4)$, $z = x + y\mathrm{i}$.

Exercise

The example and the previous exercise suggest a recipe for obtaining all homogeneous polynomials $p(x, y)$ of degree d that are harmonic. Formulate this recipe, and prove its validity.

Wave equation

The 2-dimensional wave equation is

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \qquad \text{where } a > 0 \text{ is a constant.}$$

Its solutions $u(x,t)$ describe, for example, the displacement of a vibrating string at time t and distance x from one end of the string.

Example

Show that all functions of the form

$$u(x,t) = \frac{A_0}{2} + \sum_{k=1}^{n} \Big[A_k \cos\big(k(x-at)\big) + B_k \sin\big(k(x-at)\big)\Big], \quad A_i, B_i \in \mathbb{R}$$

are solutions of the 2-dimensional wave equation.

It suffices to show this for $(x,t) \mapsto \cos\big(k(x-at)\big)$ and $(x,t) \mapsto \sin\big(k(x-at)\big)$. Let, e.g., $u(x,t) = \cos\big(k(x-at)\big)$.

$$\begin{aligned} u_x &= -k\sin(kx-kat), & u_{xx} &= -k^2\cos(kx-kat), \\ u_t &= ka\sin(kx-kat), & u_{tt} &= -k^2a^2\cos(kx-kat), \end{aligned}$$

$$\Longrightarrow u_{tt} = a^2 u_{xx}.$$

Example (cont'd)

The functions $(x,t) \mapsto \cos\bigl(k(x-at)\bigr)$, $\sin\bigl(k(x-at)\bigr)$ provide solutions of the wave equation (with given constant a^2) for all $k \in \mathbb{R}$ and can be found in a way similar to that illustrated for the Laplace equation.

The condition $k \in \mathbb{Z}$ makes these functions periodic in x (with period 2π), which is needed for a solution of the vibrating string problem.