Multivariable Calculus Reference

Sidharth Baskaran

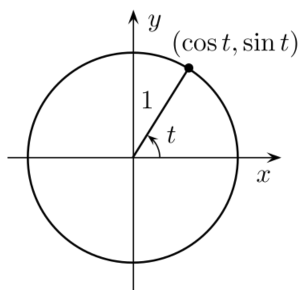
March 22, 2022

Table of Contents

1 Paths whose image curve is a circle

1.1 Unit Circle

Unit circle is set of points in $\mathbb{R}^2$ defined as $C = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 1\}$. Ellipse is $C = \{(x, y) \in \mathbb{R}^2 | \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$ Has standard parameterization of $\mathbf{c}(t) = \left(\cos(t), \sin(t)\right)$. When parameterizing, always start from $t = 0$ reference unless otherwise given.



Properties

- Image of $\mathbf{c}$ is a closed curve (has no endpoints, plane is divided into ≥ 2 disjoint regions)
- Image of $\mathbf{c}$ is a simple curve; no self-intersection
- $\mathbf{c}(t)$ is an **injective path**; path is considered injective if $\mathbf{c}(t_1) = \mathbf{c}(t_2)$, which implies that $t_1 = t_2$ where these are on the open interval (a, b) even if $a = b$
- Orientation of $\mathbf{c}$ is counter-clockwise in traversal

1.2 Observations

$$\boxed{\mathbf{p}(t) = (a\cos(\pm(nt \pm \theta)) + x_0, b\sin(\pm(nt \pm \theta)) + y_0)}$$

If $-t$ for t, orientation is CW, CCW is t. If $a = b$, then curve is a circle of radius a or b, else an ellipse with horizontal and vertical radii. x_0 and y_0 simply shift the center coordinate. $n > 0 \in \mathbb{R}$ determines how many times the circle is traversed given $t \in [0, 2\pi]$, for example. θ is the phase shift. When changing direction of traversal, cannot have $a > b$ for $[a, b]$ so to decrease argument of sin or cos must have $-t$ for t. Starting out, $-t$ goes through the angle range and t is just a sign flip.

2 Paths whose image is a line or line segment in the plane

2.1 Line Parametrics

A line is a 1D subspace of $\mathbb{R}^2$, so $L = \{t\mathbf{m} | t \in \mathbb{R}\}$ for $\mathbf{m} \in \mathbb{R}^2$. $\mathbf{m} = \begin{bmatrix} m_x \\ m_y \end{bmatrix}$ is the **slope vector**. Path given by image of L:

$$\mathbf{c}(t) = \left(m_x t, m_y t\right), t \in \mathbb{R}$$

Can represent $\mathbf{c}(t) = t\mathbf{m}$ as well.

Lines Main Ideas

- Image of a line is a curve (e.g. $y = x$ represents image curve of $\mathbf{c}(t) = (t, t)$)
- Lines can have nonzero intercepts, so $\mathbf{c}(t) = t\mathbf{m}$ represents $y = 2x + 1$. Line that has intercept vector $P_0 = (x_0, y_0) \parallel \mathbf{m} = (m_x, m_y)$ can be expressed as:

$$\boxed{\mathbf{c}(t) = (x_0 + tm_x, y_0 + m_y t) = P_0 + t\mathbf{m}}$$

  Note endpoint of $\mathbf{c}(t)$ is on image line (curve).

2.2 General Forms

2 parametric lines **collide** if they intersect and the point of intersection corresponds to the same t in both curves. If you set the parameter vector coordinates equal to each other and solve for t, a solution indicates they collide. Intersection is found by **eliminating** the parameter (solve for t in terms of either x or y and plug into the other).

General form of parameterized curve can be expressed as the following:

$$\boxed{\mathbf{c}(t) = (\frac{m_x}{\Delta t}(t-a) + x_0, \frac{m_y}{\Delta t}(t-a) + y_0)}$$

where Δt is the domain interval over $[a, b]$ and (x_0, y_0) represents the desired **starting coordinate**. This is important as when going in reverse, other coordinate can be used and slope might be negative. a is used in $(t-a)$ because everything is conventionally done with respect to starting coordinate.

3 Paths whose image curve is a line in R3

3.1 R3 parameterization

If $\mathbf{m}$ is a nonzero vector along L through origin in $\mathbb{R}^3$, then $L = \{t\mathbf{m} | t \in \mathbb{R}\}$; follows that $\mathbf{m} = (m_x, m_y, m_z)$, the slope or direction vector of the line. The basic parameterization is:

$$\mathbf{c}(t) = (m_x t, m_y t, m_z t)$$

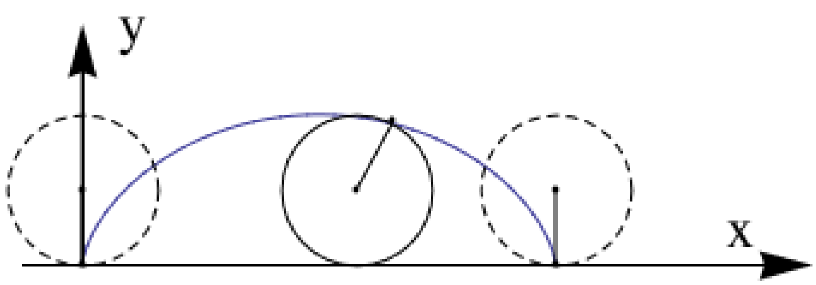
Basis vectors in $\mathbb{R}^3$ are $\mathbf{i}, \mathbf{j}, \mathbf{k}$. Rewriting parameterization:

$$\mathbf{c}(t) = (x_0 + m_x t)\mathbf{i} + (y_0 + m_y t)\mathbf{j} + (z_0 + m_z t)\mathbf{k}$$

2 lines $\mathbf{c}_1(t) = P_0 + \mathbf{m}_1 t$ and $\mathbf{c}_2(t) = Q_0 + \mathbf{m}_2 t$ are parallel if direction vectors are parallel ($m_1 = k\mathbf{m}_2$). Collisions still exist. If neither parallel nor intersecting, considered as skew.

To determine skew, parallel, or coincide, use parameters s, t for each line and solve SOE. If same slope, rule out skew clearly, then check if $s, t \in \mathbb{R}$: if not, then parallel, if so, then they coincide. If intersecting and want to check if collide, some t must satisfy all relations.

4 Cycloid Problem



With radius 1 and passing through the origin:

$$\mathbf{c}(t) = (t - \sin t, 1 - \cos t)$$

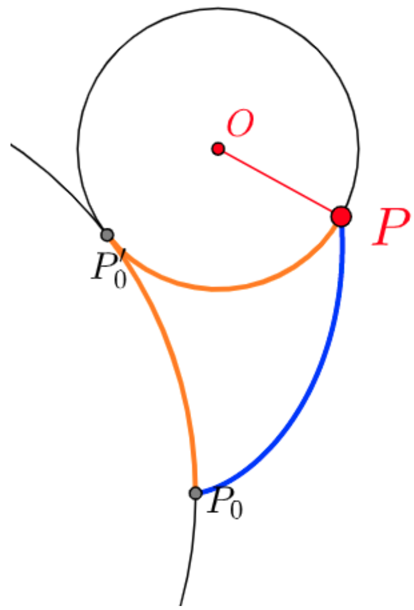
Observe that:

$$\mathbf{c}'(t) = (1 - \cos t, \sin t)$$

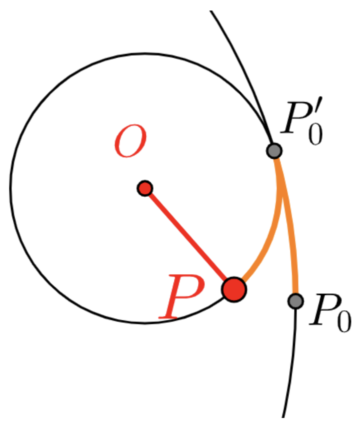
Can define the vector $\mathbf{u} = \begin{pmatrix} x'(t) \\ 0 \end{pmatrix}$ such that $\mathbf{u}$ is always horizontal and $||\mathbf{u}|| = |x'(t)|$. Reaches maximum value at $t \in [k\pi | k \in \mathbb{R}]$ and is has minimum cusp where it is 0 at $t \in [2k\pi | k \in \mathbb{R}]$. Thus, $x'(t) \geq 0$ always, as the x-coordinate is never decreasing.

Can also define the vector $\mathbf{v} = \begin{pmatrix} 0 \\ y'(t) \end{pmatrix}$ with the same properties. Reaches maximum value when $t \in [k\frac{\pi}{2} | k \in \mathbb{R}]$. Can change, as observe t when $\sin t < 0$ or > 0.

4.1 Hypercycloid Derivation



4.2 Hypocycloid Derivation



5 Velocity Vector

5.1 Definitions

Vector $\mathbf{u}(t_0) + \mathbf{v}(t_0)$ is he velocity vector to the curve $\mathbf{c}(t)$ at $t = t_0$.

Let $\mathbf{c} : [a,b] \to \mathbb{R}^n$ have a path $\mathbf{c}(t) = (x_1(t), x_2(t), x_3(t), \ldots, x_n(t))$ (let $x_i(t) : [a,b] \to \mathbb{R}$ for each i)

- If $t_0 \in [a,b]$, then $\mathbf{c}'(t_0) := \left(x_1'(t_0), x_2'(t_0), x_3'(t_0), \ldots, x_n'(t_0)\right)$; the velocity vector to $\mathbf{c}$ at t_0
- The path $\mathbf{c}'(t_0) := \left(x_1'(t_0), x_2'(t_0), x_3'(t_0), \ldots, x_n'(t_0)\right)$; the velocity vector to $\mathbf{c}$ is referred to as

velocity of $\mathbf{c}(t)$

Recall chain rule: if $y = f(x)$ where x is a function of t, $y'(t) = x'f'(x)$, not to be confused with product rule. Can write $f'(x) = \frac{y'(t)}{x'(t)}$

- If $\mathbf{p}(t) = \mathbf{c}(t) + \mathbf{r}(t)$, then $\mathbf{p}'(t) = \mathbf{c}'(t) + \mathbf{r}'(t)$
- If $g(t) = \mathbf{c}(t) \cdot \mathbf{r}(t)$, then $g'(t) = \mathbf{c}'(t) \cdot \mathbf{r}(t) + \mathbf{c}(t) \cdot \mathbf{r}'(t)$
- If $\mathbf{p}(t) = f(t)\mathbf{c}(t)$, then $\mathbf{p}'(t) = f'(t)\mathbf{c}(t) + f(t)\mathbf{c}'(t)$
- If $\mathbf{p}(t) = \mathbf{c}(t) \times \mathbf{r}(t)$, then $\mathbf{p}'(t) = \mathbf{c}'(t) \times \mathbf{r}(t) + \mathbf{c}(t) \times \mathbf{r}'(t)$
- If $\mathbf{p}(t) = \mathbf{c}(f(t))$, then $\mathbf{p}'(t) = f'(t)\mathbf{c}'(f(t))$
- If $g(t) = \|\mathbf{c}(t)\|$, then $g'(t) = \frac{\mathbf{c}(t) \cdot \mathbf{c}'(t)}{\|\mathbf{c}(t)\|}$

5.2 Tangent Line

Tangent line can be visualized as a base vector in standard position plus a velocity vector tangent to the tip which traces a shifted line in some interval. General formula with base vector $\mathbf{c}(t_0)$ and slope $\mathbf{c}'(t_0)$:

$$\boxed{\ell(t) = \mathbf{c}(t_0) + (t - t_0)\mathbf{c}'(t_0)}$$

6 Space Curves

- Projection into the xy plane is the path $(x(t), y(t), 0)$.
- Projection into the $xz-$ plane is the path $(x(t), 0, z(t))$.
- Projection into the yz plane is the path $(0, y(t), z(t))$.

7 Speed and Arclength

7.1 Speed

Speed of a parametric function in $\mathbb{R}^n$ is given by:

$$||\mathbf{c}'(t)|| = \sqrt{\sum_{i=1}^{n} c_i(t)^2}$$

(being the magnitude of the velocity vector)

7.2 Arclength

Arclength of a parametric function is given by:

$$S = \int_a^b ||\mathbf{c}'(t)||dt = \int_a^b \sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2 + (\frac{dz}{dt})^2 + \cdots}\, dt$$

Can approximate arclength as a sum of the lengths of secant vector approximations $\mathbf{s}_i = \mathbf{c}(t_i) - \mathbf{c}(t_{i-1})$:

$$\text{arclength} \approx \sum_{i=1}^{n} ||\mathbf{s}_i||$$

According to the MVT, there exists a $\hat{t}_i$ in (t_{i-1}, t_i) (open interval due to differentiability requirement) such that:

$$x'(\hat{t}_i) = \frac{x(t_i) - x(t_{i-1})}{t_i - t_{i-1}}$$
$$y'(\hat{t}_i) = \frac{y(t_i) - y(t_{i-1})}{t_i - t_{i-1}}$$

This means that, since $\mathbf{s}_i$ is given as the difference between 2 points, being a secant:

$$\mathbf{s}_i = \big((t_i - t_{i-1})x(\hat{t}_i), (t_i - t_{i-1})y(\hat{t}_i)\big)$$
$$\mathbf{s}_i = (t_i - t_{i-1})\big(x'(\hat{t}_i), y'(\hat{t}_i)\big)$$
$$\mathbf{s}_i = (t_i - t_{i-1})\mathbf{c}'(\hat{t}_i)$$

Thus,

$$\text{arclength} \approx \sum_{i=1}^{n} ||\mathbf{s}_i||$$
$$\text{arclength} \approx \sum_{i=1}^{n} ||\Delta t\, \mathbf{c}'(\hat{t}_i)||$$
$$\text{arclength} \approx \sum_{i=1}^{n} \Delta t ||\mathbf{c}'(\hat{t}_i)||$$

Can define the arclength differential as follows:

$$\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$$

Can just define arclength as arclength $= \int \mathrm{d}s$

7.3 Arclength Parameterization

Higher the speed of a curve, farther the points are spaced apart. An arclength parametrization of a curve is a path whose image is the desired curve and whose speed is constantly one. Or, $\mathbf{c} : [a,b] \to \mathbb{R}^n$ with $||\mathbf{c}'(t)|| = 1$ for $t \in [a,b]$. If a curve is not an arclength parameterization, then can do $\frac{\mathbf{c}(t)}{||\mathbf{c}'(t)||}$ but only dividing the coefficients (slopes).

When speed is variable, is difficult to define arclength parameterization. Thus, can define displacement to be $s(t) = \int_a^b \mathbf{c}(t)dt$. If $v(t)! \neq 0$, then s is injective because according to FTC, $s'(t) = v(t)$. By definition, $v'(t) \geq 0$ always since it is composed of a radical, so it must be **increasing**. Thus, if $t_1 = t_2$, $s(t_1) \neq s(t_2)$. Arclength parameterization:

$$s(t) = \int_0^t ||\mathbf{c}'(u)||du$$

This means that s is invertible, so can solve for t to get $t = \varphi(s)$. An arclength parameterization can be found by:

$$\boxed{\mathbf{p}(s) = \mathbf{c}(\varphi(s))}$$

8 Curvature

8.1 Proofs

Recall that to make an arclength parameterization accumulate the magnitudes of infinitesimal velocity vectors:

$$s(t) = \int_a^t ||\mathbf{r}'(u)||du$$

Given some curve $\mathbf{r}(t)$, define an arclength parameterization by $\mathbf{r}(g(s)) \to \mathbf{r}_1(s)$, so $\mathbf{r}$ is defined in terms of s. The unit tangent vector $\mathbf{T}_1(s)$ is then $\frac{\mathbf{r}_1{}'(s)}{||\mathbf{r}_1{}'(s)||} = \mathbf{r}_1{}'(s)$.

$$\begin{aligned}
\mathbf{T}_1(s) &= \mathbf{r}_1{}'(s) \\
&= \frac{d}{ds}\mathbf{r}_1(s) \\
&= \frac{d}{ds}\mathbf{r}(g(s)) \\
&= \mathbf{r}'(g(s)) \cdot g'(s) = \mathbf{r}'(t) \cdot \frac{dt}{ds} \\
&= \frac{\mathbf{r}'(t)}{\frac{ds}{dt}} = \frac{\mathbf{r}'(t)}{||\mathbf{r}'(t)||}
\end{aligned}$$

This means that $\mathbf{T}_1(s) = \mathbf{T}(t)$

Continuing, to find curvature $\kappa(t)$:

$$\begin{aligned}
\mathbf{T}_1{}'(s) &= \frac{d}{ds}\mathbf{T}(t) \\
&= \frac{d}{ds}\mathbf{T}(g(s)) \\
&= \mathbf{T}'(t)\cdot\frac{dt}{ds} \\
&= \frac{\mathbf{T}'(t)}{\frac{ds}{dt}} \\
&= \frac{\mathbf{T}'(t)}{||\mathbf{r}'(t)||}
\end{aligned}$$

Thus, $\boxed{\kappa(t) = \dfrac{||\mathbf{T}'(t)||}{||\mathbf{r}'(t)||}}$

8.2 Definition

Given a curve C parameterized with arclength by the path $\mathbf{c} : [a,b] \to \mathbb{R}^n$, curvature is defined as:

$$\boxed{\kappa(s) = ||\mathbf{T}'(s)||}$$

where $\mathbf{c}'(s) \neq 0$ and $\mathbf{T}(s) = \frac{\mathbf{c}'(s)}{||\mathbf{c}'(s)||}$ (normalized slope vector).

A loose geometric interpretation is that a greater $\kappa(s)$ implies more curvature, that is, the curve is changing at a greater rate there. When $\mathbf{c}'(s) \neq 0$ is always true for a curve, it is **regular**. Is defined in terms of arclength parameterization so curvature is an intrinsic property of the curve independent of parameterization.

Formula for curvature at the point $\mathbf{c}(t)$:

$$\boxed{\kappa(t) = \frac{||T'(t)||}{||\mathbf{c}'(t)||} = \frac{||\mathbf{c}'(t) \times \mathbf{c}''(t)||}{||\mathbf{c}'(t)||^3}}$$

9 Motion in 3D space

Given a path $\mathbf{c} : \mathbb{R} \to \mathbb{R}^3$ with $\mathbf{c}(t) = (x(t), y(t), z(t))$, then we have defined:

- $\cdot\mathbf{v}(t) = \mathbf{c}'(t) = \left(x'(t), y'(t), z'(t)\right)$ is also path in $\mathbb{R}^3$ called the velocity of $\mathbf{c}$
- $\mathbf{a}(t) = \mathbf{v}'(t) = \mathbf{c}''(t) = \left(x''(t), y''(t), z''(t)\right)$ is also a path in $\mathbb{R}^3$ called the acceleration of $\mathbf{c}$

- $v(t) = \|\mathbf{v}(t)\| = \left\|\mathbf{c}'(t)\right\|$ is a scalar valued function on $\mathbb{R}$ (that's a fancy way of saying the domain and codomain of this function are both $\mathbb{R}$) called the speed of $\mathbf{c}$
- $\mathbf{T}(t) = \frac{\mathbf{c}'(t)}{v(t)}$ is also a path in $\mathbb{R}^3$ called the unit tangent to $\mathbf{c}$
- $\kappa(t) = \frac{\mathbf{T}'(t)}{v(t)} = \frac{\|\mathbf{c}'(t)\times\mathbf{c}''(t)\|}{\|\mathbf{c}'(t)\|^3}$ is a scalar valued function on $\mathbb{R}$ called the curvature of $\mathbf{c}$

Note that $\mathbf{T}\cdot\mathbf{T} = ||\mathbf{v}||^2 = 1$. Computing the derivative, $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{T}\cdot\mathbf{T} = 2\mathbf{T}\cdot\mathbf{T}' = 0$ This means that $\mathbf{T}\perp\mathbf{T}'$.

Define $\boxed{\mathbf{N}(t) = \dfrac{\mathbf{T}'(t)}{\left\|\mathbf{T}'(t)\right\|}}$ as the unit normal vector, which is the unit tangent to the unit tangent. From observation, $\mathbf{T}\perp\mathbf{N}$. Fact: the acceleration vector always lies in the plane spanned by $\mathbf{N}$ and $\mathbf{T}$.

Acceleration $\mathbf{a}(t)$ is thus split component-wise into a_T from $\mathbf{T}$ and a_N from $\mathbf{N}$:

$$\boxed{a_T = v'(t) = \frac{\mathbf{a}(t)\cdot\mathbf{v}(t)}{v(t)}}$$

$$\boxed{a_N = \kappa(t)v(t)^2 = \frac{||\mathbf{a}(t)\times\mathbf{v}(t)||}{v(t)} = \sqrt{\|\mathbf{a}(t)\|^2 - |a_T|^2}}$$

10 Derivatives of parameterized curves

10.1 Arclength parameterization derivation

Take the following function:

$$\mathbf{r}(t) = \langle x(t), y(t)\rangle$$

An arclength parameterization is achieved with the following computation:

$$s = \int_0^t \sqrt{x'(u)^2 + y'(u)^2}\,du$$

Can say that $t = g(s)$, so the arclength parameterization, which is the path in terms of s:

$$\mathbf{r}_1(s) = \langle x(g(s)), y(g(s))\rangle$$

Taking the derivative by the chain rule:

$$\begin{aligned}\mathbf{r}_1{}'(s) &= \langle x'(g(s))\cdot g'(s), y'(g(s))\cdot g'(s)\rangle \\ &= g'(s)\langle x'(t), y'(t)\rangle\end{aligned}$$

Note that $g'(s) = \frac{dt}{ds} = \frac{1}{\frac{ds}{dt}} = \frac{1}{||\mathbf{r}'(t)||}$ by taking the derivative of the integral for arclength:

$$\mathbf{r}_1{}'(s) = \frac{1}{||\mathbf{r}'(t)||}\langle x'(t), y'(t)\rangle = \mathbf{T}(t)$$

Following from this, $g'(s) = \frac{1}{||\mathbf{r}'(g(s))||}$ so $g''(s) = -\frac{1}{||\mathbf{r}'(g(s))||^2} \cdot g'(s) = -\frac{1}{||\mathbf{r}'(t)||^3}$

10.2 Orthogonal derivative and position vectors

Observe that $\mathbf{r}\cdot\mathbf{r} = ||\mathbf{r}||^2$. Thus, $\frac{d}{dt}[\mathbf{r}\cdot\mathbf{r}] = 2\mathbf{r}\cdot\mathbf{r}' = 2||\mathbf{r}||||\mathbf{r}||'$. Rearranging: $\frac{\mathbf{r}\cdot\mathbf{r}}{||\mathbf{r}||} = ||\mathbf{r}||'$. Means that magnitude of position vector has to be a constant value in order for it to be $\perp$ to derivative.

11 Planetary motion

- Law of ellipses – orbit of planet is ellipse with sun as focus
- Law of equal area in equal time – position vector pointing from sun to planet sweeps out equal area in equal time (so speed must increase/decrease)

Can approximate the area swept in time by $\frac{dA}{dt} = \frac{1}{2}||\mathbf{r}(t)\times\mathbf{r}'(t)|| = \frac{1}{2}||\mathbf{J}||$. The differential equation for each of Kepler's laws is: $\mathbf{r}''(t) = -\frac{k}{||\mathbf{r}(t)||^3}\mathbf{r}(t)$, so it is in the direction of $\mathbf{r}(t)$. Thus, differentiating $\frac{d\mathbf{J}}{dt} = \frac{d}{dt}(\mathbf{r}'(t)\times\mathbf{r}''(t)) = 0$.

11.1 Cross-product identities

Cross product identities:

- $\boxed{\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = (u\cdot w)\mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\mathbf{w}}$
- $\boxed{u\cdot(\mathbf{v}\times\mathbf{w}) = v\cdot(\mathbf{w}\times\mathbf{u}) = \mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})}$

12 Planes

If (x,y,z) is a point on the plane, then given $\mathbf{P}_0 = (x_0, y_0, z_0)$, $(x-x_0, y-y_0, z-z_0)$ is a vector on the plane perpendicular to $\mathbf{n}$, the normal vector. Thus, $(A,B,C)\cdot(x-x_0, y-y_0, z-z_0) = 0$ where A,B,C are vector coordinates of $\mathbf{n}$. With expansion:

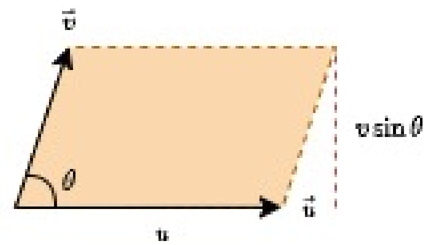
$$\begin{aligned} A(x-x_0)+B(y-y_0)+C(z-z_0) &= 0 \\ Ax+By+Cz = Ax_0+By_0+Cz_0 &= 0 \\ Ax+By+Cz &= \mathbf{n}\cdot\mathbf{P}_0 \end{aligned}$$

Note that (A,B,C) form coordinates of $\mathbf{n}$.

To find a plane containing 3 points $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, compute, for example $\mathbf{c}_1 = \mathbf{v}_3 - \mathbf{v}_1$ and $\mathbf{c}_2 = \mathbf{v}_2 - \mathbf{v}_1$. This finds 2 vectors in the plane. Then compute $\mathbf{c}_1\times\mathbf{c}_2 = \mathbf{n}$.

The trace of a plane is the intersection of a plane $\mathscr{P}$ with xy, xz, or yz coordinate planes. Can be found by setting respective variable to 0.

12.1 Cross-product rules and identities



Overview

- $||\mathbf{a} \times \mathbf{b}|| = ||\mathbf{a}||||\mathbf{b}|| \sin\theta$
- $\mathbf{a}, \mathbf{b} \perp \mathbf{a} \times \mathbf{b}$

Algebraic

- $\mathbf{a} \times \mathbf{b} = \mathbf{0}$
- $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
- Distributive properties hold – preserve direction however
- $(\alpha\mathbf{a}) \times \mathbf{b} = \alpha(\mathbf{a} \times \mathbf{b})$

13 Graphs

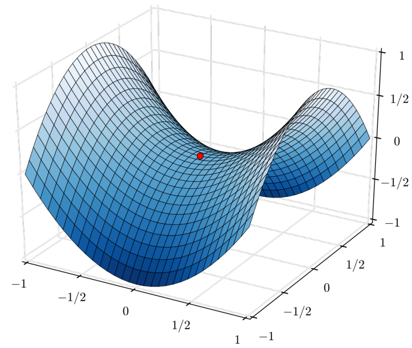
13.1 Multivariable functions

Function of n-variables is real-valued function with $f(x_1, \cdots, x_n)$ with domain $\mathscr{D}$ being a set of n-tuples $(x_1, \cdots, x_n)$ in $\mathbb{R}^n$, or where f is defined. Range of f is all values $f(x_1, \cdots, x_n)$ for $(x_1, \cdots, x_n)$ in the domain.

13.2 Graphing multivariable functions

Traces are 2D curves obtained by intersection with planes parallel to coordinate plane.

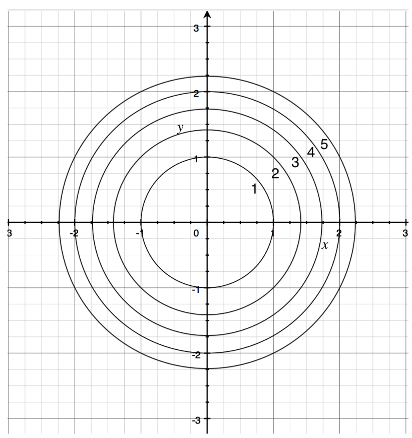
- Horizontal trace at height c – intersection of graph with plane $z = c$, so points (x, y, c) such that $f(x, y) = c$
- Vertical trace in plane $x = a$ – intersection of graph with vertical plane $x = a$ for all points $(a, y, f(a, y))$
- Vertical trace in plane $y = b$ – intersection of graph with vertical plane $y = b$ for all points $(x, b, f(x, b))$



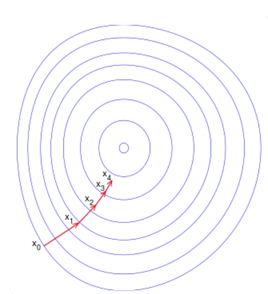
Saddle surface general form is $f(x,y) = x^2 - y^2$. The horizontal traces are hyperbolas of the form $c = x^2 - y^2$. Vertical traces are parabolas, as either x, y set to 0.

Linear functions in 2 variables are of the form $f(x,y) = mx + ny + r | m, n, r \in \mathbb{R}$.

13.3 Contour maps and level curves



Can specify a contour interval for each $z = c$ value. Is a 2D representation of level curves of $f(x,y)$ at an interval. Going along level curve means change in altitude is 0. Altitude has change of $\pm m$ (contour interval) when going up/down contour levels. Average ROC is Δelevation/Δdistance. Path of steepest ascent follows the shortest possible segment from one contour line to another and always points in steepest direction.



14 Partial Derivatives

14.1 Definition

If $f : \mathbb{R}^2 \to \mathbb{R}$ is given by $f(x, y) = z$ abd $P_0 = (a, b)$ is a point in the domain of f, then the partial derivative are:

- If $h : \mathbb{R} \to \mathbb{R}$ by $h(t) = f(t, b)$, then partial derivative with respect to x at P_0 is $h'(a)$ with following limit definition

$$\left.\frac{\partial f}{\partial x}\right|_{(a,b)} = f_x(a, b) = \lim_{h\to 0} \frac{f(x+h, b) - f(x, b)}{h} = \lim_{x\to a} \frac{f(x, b) - f(a, b)}{x - a}$$

- If $g : \mathbb{R} \to \mathbb{R}$ by $g(t) = g(t, b)$ then partial derivative with respect to y at P_0 is $g'(b)$ with following limit definition

$$\left.\frac{\partial f}{\partial y}\right|_{(a,b)} = f_y(a, b) = \lim_{h\to 0} \frac{f(a, x+h) - f(a, b)}{h} = \lim_{y\to b} \frac{f(a, y) - f(a, b)}{y - b}$$

Can be thought of as the intersection of the plane shifted by b with f, and the derivative of the resulting trace.

14.2 Linear approximation with planes

Let $z = f(x, y)$ be a scalar-valued function in $\mathbb{R}^2$ and $P_0 = (a, b)$ be a point in domain of f. Can have 2 slope vectors representing partial derivatives: $(1, 0, f_x(a, b))$ and $(0, 1, f_y(a, b))$. Can find a linear approximation by finding set of points in plane spanned by these vectors passing through $(a, b, f(a, b))$.

$$\mathbf{n} = (1, 0, f_x(a, b)) \times (0, 1, f_y(a, b)) = (-f_x(a, b), -f_y(a, b), 1)$$

Building the plane:

$$\begin{aligned}
(x - a, y - b, z - f(a, b)) \cdot \mathbf{n} &= 0 \\
(x - a, y - b, z - f(a, b)) \cdot \left(-f_x(a, b), -f_y(a, b), 1\right) &= 0 \\
-f_x(a, b)(x - a) - f_y(a, b)(y - b) + z - f(a, b) &= 0
\end{aligned}$$

Thus,

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

14.3 Higher-order derivatives

Can be calculated using derivatives of f_x and f_y. Notation:

$$f_{xx} = \frac{\partial}{\partial x}(\frac{\partial f}{\partial x}),\ f_{yy} = \frac{\partial}{\partial y}(\frac{\partial f}{\partial y})$$

Can also have mixed partials (read as with respect to x or y):

$$f_{xy} = \frac{\partial}{\partial y}(\frac{\partial f}{\partial x}),\ f_{yx} = \frac{\partial}{\partial x}(\frac{\partial f}{\partial y})$$

By Clairaut's Theorem, if f_{xy} and f_{yx} are both continuous functions on a disk D, then $f_{xy}(a,b) = f_{yx}(a,b)\,\forall (a,b) \in D$. Means that $f_{xyxy} = f_{xxyy} = f_{yyxx} = f_{yxyx}$.

15 Extrema

15.1 Definition and proofs

A function f has:

- local maximum at P_0 in domain if $f(P_0) > f(x,y)\ \forall\ (x,y)$ sufficiently near P_0
- local minimum at P_0 in domain if $f(P_0) < f(x,y)\ \forall\ (x,y)$ sufficiently near P_0

Sufficiently near: positive radius R used to build a circle centered at P_0 that traps points in domain with desired property (min or max). Global extrema redefine sufficiently near as in the domain of f.

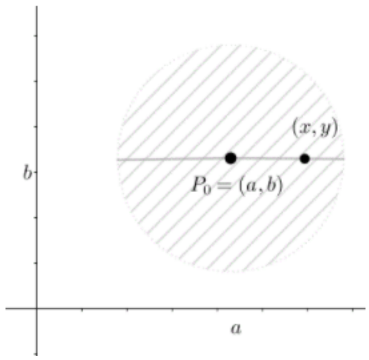
Critical point is defined as either of the following:

- $f_x(P_0) = f_y(P_0) = 0$
- either $f_x(P), f_y(P)$ does not exist

Method: find critical value y from f_y and x from f_x through cross-substitution.

Proof that if $f_x(P), f_y(P)$ both exist and there is a local max at P, then both partials are 0:

Define $g : \mathbb{R} \to \mathbb{R}$ to be single variable function from holding $y = b$ in f. Considering points sufficiently near:



Plugging into function g: $g(x) < f(x,b) \leq f(P_0) = g(a)$, so for x-values sufficiently near a, $g(x) \leq g(a)$, so it is a local max of g. Invoking theorem of extrema, either $g'(a) = 0$ or DNE. As $g'(a) = f_x(P_0)$, demonstrates that $f_x = 0$ or DNE for a local max. Same can be done for f_y.

Can minimize a function's distance to origin through distance formula. As a square root minimizes at its argument, can just concentrate on minimizing the argument.

Reiteration: If $f : \mathbb{R}^2 \to \mathbb{R}$ has continuous first and second order partial derivatives then $f_{xy}(x,y) = f_{yx}(x,y)$. Continuous 2nd order partials: C^2; continous n order partials: C^∞.

Fermat's theorem proof (same as above): If $f(x,y)$ has a local min/max at $P = (a,b)$, then $P = (a,b)$ is a critical point of $f(x,y)$.

Assuming $f(x,y)$ has a local min at P, then this means that $f(x,y) \geq (a,b)$ for (x,y) in the surrounding disk $D(r,P)$. For some $y = b$, the distance between any 2 x-values must be contained in the disk: $|x - a| < r$. Shows that $g(x) = f(x,b)$ has a local min at $x = a$, so $g'(a) = 0$ or DNE. Because $g'(a) = f_x(a,b)$, $f_x(a,b)$ is either 0 or DNE. As the same can be said for f_y, P is a critical point.

15.2 Second-derivative test

Given a function $f : \mathbb{R}^2 \to \mathbb{R}$ that has continuous 2nd partials near a critical point P, define the discriminant of f at P to be:

$$D(P) = f_{xx}(P) f_{yy}(P) - \left(f_{xy}(P)\right)^2$$

Then:

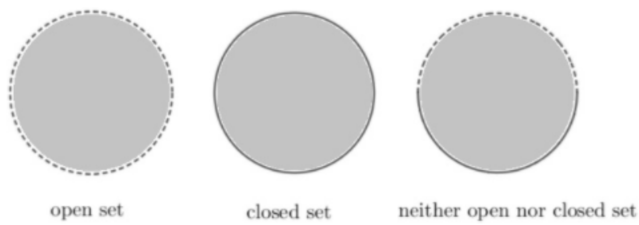
- If $D > 0$, P is a local extreme of f
  - $f_{xx}(P) > 0$ implies a local min
  - $f_{xx}(P) < 0$ implies a local max
- If $D < 0$ then there is neither a min or max at P, but an inflection point (saddle point)
- If $D = 0$ then there is no information about P
- If $D > 0$ and $f_{xx}(a,b) > 0$, then local minimum
- If $D > 0$ and $f_{xx}(a,b) < 0$, then local maximum

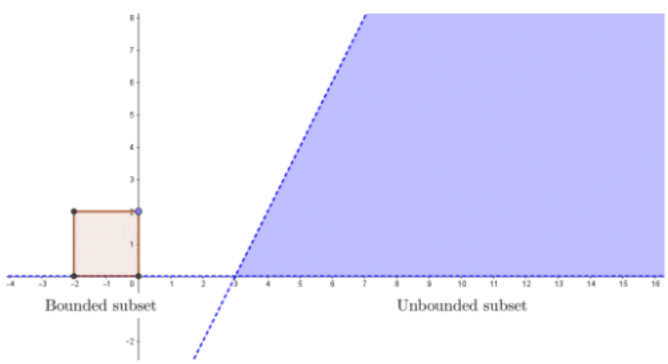
Observe that $D(P) = \det\begin{pmatrix} f_{xx}(P) & f_{xy}(P) \\ f_{yx}(P) & f_{yy}(P) \end{pmatrix}$

15.3 Global extrema

If $f : \mathbb{R}^2 \to \mathbb{R}$ is continuous on a closed and bounded subset of $\mathbb{R}^2$ then it has a global max/min on the subset (at critical point or along boundary)

A closed subset of $\mathbb{R}^2$ is one that contains the boundary. Boundary points have the property that any circle centered around them with positive radius will contain points in and out of the subset. Bounded subset is where any distance between 2 points in the set never exceeds some fixed bound $M \in \mathbb{R}$.





The boundary curve is denoted as ∂D where D is a disk. By parameterizing ∂D and using function composition to take this curve $\mathbf{c}(t)$ into f: $h(t) = f(\mathbf{c}(t))$. A min/max for f along boundary curve is a min/max of h. Plug resulting coordinates into f and determine global extrema.

16 Lagrange Multipliers

16.1 Theory

Let $f : \mathbb{R}^2 \to \mathbb{R}$ and $\mathbf{c} : \mathbb{R} \to \mathbb{R}^2$ such that there is a composite function $h(t) = f(\mathbf{c}(t))$. Then, $h'(t) = f_x(\mathbf{c}(t))x'(t) + f_y(\mathbf{c}(t))y'(t) = \left(f_x\left(\mathbf{c}(t)\right), f_y\mathbf{c}(t)\right) \cdot \mathbf{c}'(t)$. This brings definition of gradient vector, so that for $f : \mathbb{R}^2 \to \mathbb{R}$, each point in domain is in domain of f the vector $\left(f_x(P_0), f_y(P_0)\right)$:

$$\nabla f(P_0) = \left(f_x(P_0), f_y(P_0)\right)$$

Rewriting chain rule:

$$\frac{d}{dt}f(\mathbf{c}(t_0)) = \nabla f(\mathbf{c}(t_0)) \cdot \mathbf{c}'(t_0)$$

Let $g(x,y)$ be boundary curve function and $f(x,y)$ be original.

Use a boundary curve which is level curve $c = 0$ which represents the boundary of the subset where global extrema can be found. If f achieves maximum at point P along curve, then either $g_x(P) = g_y(P) = 0$ ($\nabla g(P) = \mathbf{0}$) or there is a scalar λ (Lagrange multiplier) such that $f_x(P) = \lambda g_x(P)$ and $f_y(P) = \lambda g_y(P)$ ($\nabla f(P) = \lambda \nabla g(P)$). Can find locations of global extreme without parameterizing.

16.2 Proof

Suppose $f(P) \geq f(x,y)\ \forall\ (x,y)$ satisfying $g(x,y) = 0$ where g is the bounded constraint. Let $\mathbf{c}$ be its parameterization where $\mathbf{c}'(0) \neq 0$. Parameterizing $g(x,y) = 0$ with $\mathbf{c}(t)$ such that $\mathbf{c}(0) = P$: Observe that $h(t) = f(\mathbf{c}(t))$ is from $\mathbb{R} \to \mathbb{R}$ with local max at $t = 0$. Thus, $h'(t) = 0$ or DNE. Taking derivative, $h'(t) = \nabla f(\mathbf{c}(t)) \cdot \mathbf{c}'(t)$. Then, $h'(0) = \nabla f(\mathbf{c}(0)) \cdot \mathbf{c}'(0) = 0$ or DNE. This dot product is either 0 or DNE because 1 or both vectors could not exist, or if they are $\perp$. Thus, $\nabla f(P) \perp \mathbf{c}'(0)$ and $\mathbf{c}'(0) \neq \mathbf{0}$ (because $\mathbf{c}$ is assumed to be regular).

Showing $\nabla g(P) \perp \mathbf{c}'(0)$. Define $j(t) = g(\mathbf{c}(t))$. Thus, $j'(t) = \nabla g(\mathbf{c}(t)) \cdot \mathbf{c}'(t)$. Because $\mathbf{c}(t)$ is a parameterization of level curve $g(x,y) = 0$, it always outputs 0, so $j(t) = g(\mathbf{c}(t)) = 0$ always. Can the conclude that $j'(t) = \nabla g(\mathbf{c}(t)) \cdot \mathbf{c}'(t) = 0$. Plugging in $t = 0$ gives $\nabla g(\mathbf{c}(0)) \cdot \mathbf{c}'(0) = 0$ so $\nabla g(P) \perp \mathbf{c}'(0)$.

Because both vectors are perpendicular to $\mathbf{c}'(0)$, they are parallel – $\nabla f(P) = \lambda \nabla g(P)$. Or:

$$f_x(P) = \lambda g_x(P)$$
$$f_y(P) = \lambda g_y(P)$$

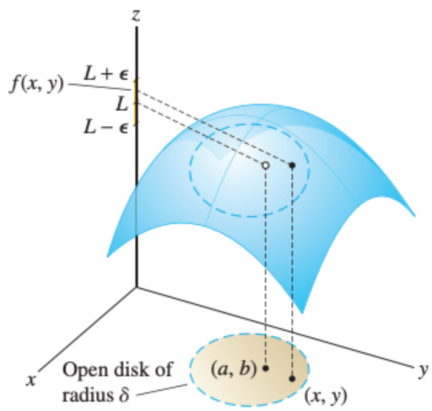
17 Limits

17.1 Delta-Epsilon Definition on R

The limit of f as x approaches a is L if given a positive real number ϵ, there exists a corresponding δ so that numbers selected with distance from a less than δ and > 0 it is guaranteed that outputs under f are within distance ϵ from L.

Formal definition

$\lim_{x \to a} f(x) = L$ if $\forall \epsilon > 0$, $\exists \delta > 0$ such that $0 < |x - a| < \delta \implies |f(x) - L| < \epsilon$



17.2 Definition of Limit on R2

If $f : \mathbb{R}^2 \to \mathbb{R}$ and $P_0 = (x_0, y_0)$, then $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L$ if given any small positive number ϵ, $\delta > 0$ is guaranteed so when there is a point $C \neq P_0$ within circle of radius δ about P_0, $f(x,y)$ lies within ϵ of L.

Formal definition

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = \lim_{\mathbf{x}\to P_0} f(x,y) = L \text{ if } \forall\varepsilon > 0\, \exists\delta > 0$$

such that

$$0 < \sqrt{(x-x_0)^2 + (y-y_0)^2} < \delta \implies |f(x,y) - L| < \varepsilon$$

Continuous functions

- Polynomials: $f(x,y) = x^2y + xy^2$
- Exponentials: $f(x,y) = e^{xy}$
- Trigonometric: $f(x,y) = \sin(x+y)$
- Compositions of any continuous functions: $f(x,y) = \cos(e^{x^2y - xy^2})$
- Sums, differences, products of continuous functions: $f(x,y) = x^2y - xy^2 + e^{xy}\sin(xy)$
- Quotients of continuous functions (domain of quotient does not include zeros of denominator): $f(x,y) = \frac{x^{2y} - xy^2}{\sin(xy)}$

17.3 Computational Techniques

Can use continuity to find limit (plug in point). Can also use conjugate multiplication to simplify and compute. Cancelling terms works because a limit approaches a value instead of equaling it. Can also squeeze a function between 2 continuous ones to find limit.

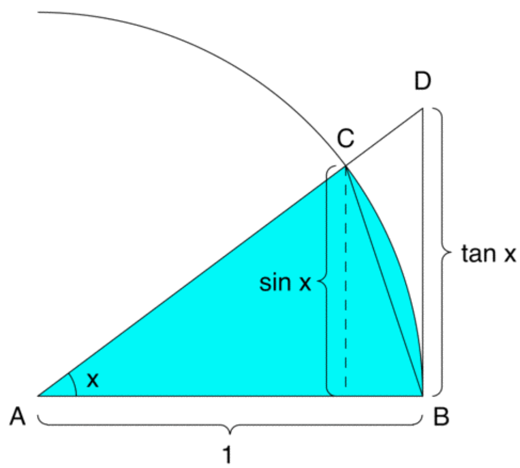
Let $\lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}}$ exist. Know that $0 \le \lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}}$. Because $y^2 > 0$, $\frac{x^2}{\sqrt{x^2+y^2}} \le \frac{x^2+y^2}{\sqrt{x^2+y^2}}$. Thus, $\lim_{(x,y)\to(0,0)} 0 \le \lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}} \le \lim_{(x,y)\to(0,0)} \left(x^2+y^2\right)^{1/2}$ so the limit is 0.

Other strategies: If a term $k(x,y) \geq 0 \; \forall \; (x,y) \in \mathbb{R}^2$, then can simply perform ± 1 to denominator for squeeze theorem proofs.

Common inequalities

- AM-GM inequality: $\frac{x+y}{2} \geq \sqrt{xy} \implies (x-y)^2 \geq 0$
- Triangle inequality: $|x+y| \leq |x| + |y|$ and $|x-y| \geq |x| - |y|$
- $|e^x - 1| \leq |x| e^{|x|}$
- $(x \pm y)^2 \geq 0$

Proof that $\frac{\sin x}{x} = 1$:



It is evident that:

$$\frac{1}{2}\sin x \leq \frac{1}{2}x \leq \frac{1}{2}\tan x \tag{1}$$

$$\sin x \leq x \leq \tan x \tag{2}$$

$$1 \leq \frac{x}{\sin x} \leq \frac{1}{\cos x} \tag{3}$$

$$1 \geq \frac{\sin x}{x} \geq \cos x \tag{4}$$

Applying the squeeze theorem to 4:

$$\lim_{x \to 0} 1 \geq \lim_{x \to 0} \frac{\sin x}{x} \geq \lim_{x \to 0} \cos x$$

$$1 \geq \frac{\sin x}{x} \geq 1$$

Thus, $\lim_{x \to 0} \frac{\sin x}{x} = 1$.

18 Derivatives

18.1 Limit Definition

A function $f:\mathbb{R}^m \to \mathbb{R}^n$ is differentiable at P_0 if there is a linear transformation from $n \times m$ matrix $Df(P_0)$ satisfying

$$\lim_{\mathbf{x}\to P_0} \frac{||f(\mathbf{x}) - \left(f(P_0) + \mathrm{D}f(P_0)(\mathbf{x}-P_0)\right)||}{||\mathbf{x}-P_0||} = 0$$

If the function is differentiable at this point, then $Df(P_0)$ is the derivative of f at P_0. Called Jacobian matrix of f at P_0.

Can adapt this to single-variable calculus:

$$\lim_{x\to a} \frac{|f(x) - \left(f(a) + (m_a)(x-a)\right)|}{|x-a|} = 0$$

Simplifying and reconfiguring to limit definition of a derivative:

$$\begin{aligned}
\lim_{x\to a} \frac{|f(x) - \left(f(a) + (m_a)(x-a)\right)|}{|x-a|} &= 0\\
\lim_{x\to a} \left|\frac{f(x) - \left(f(a) + (m_a)(x-a)\right)}{x-a}\right| &= 0\\
\left|\lim_{x\to a} \frac{f(x) - \left(f(a) + (m_a)(x-a)\right)}{x-a}\right| &= 0\\
\lim_{x\to a} \frac{f(x) - \left(f(a) + (m_a)(x-a)\right)}{x-a} &= 0 \,\text{since}\, |0| = 0\\
\lim_{x\to a} \left(\frac{f(x)-f(a)}{x-a} - m_a\right) &= 0\\
\lim_{x\to a} \frac{f(x)-f(a)}{x-a} &= m_a
\end{aligned}$$

Equivalent to the case of a 1×1 matrix $Df(P_0)$ where the single entry is $f'(a)$. Also $m_a = f'(a) \implies f(x) = f(a) + f'(a)(x-a)$ is the linear approximation.

As x-values approach a, $f(x) - (f(a) + f'(a)(x-a))$ approaches 0 faster. Thus, $\frac{f(x)-\left(f(a)+f'(a)(x-a)\right)}{x-a}$ approaches 0.

18.2 Multivariable Application

Approximating plane of some $f:\mathbb{R}^2 \to \mathbb{R}$ is $L_{P_0}(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$, i.e. the plane spanned by $\langle 1, 0, f_x(a,b)\rangle$ and $\langle 0, 1, f_y(a,b)\rangle$ passing through $(a, b, f(a,b))$.

Is identical to $\begin{bmatrix} f_x(P_0) & f_y(P_0)\end{bmatrix}\begin{bmatrix} x-a \\ y-b\end{bmatrix}$. The matrix is the matrix of partial derivatives.

Such a function is differentiable if:

$$\lim_{(x,y)\to(a,b)} \frac{f(x,y)-\left(f(a,b)+\begin{pmatrix}f_x(a,b) & f_y(a,b)\end{pmatrix}\begin{pmatrix}x-a\\ y-b\end{pmatrix}\right)}{\sqrt{(x-a)^2+(y-b)^2}} = 0$$

Numerator is linear approximation and denominator is distance to point (a,b). Thus:

Differentiability

$f(x,y)$ is differentiable at the point P_0 if $\displaystyle\lim_{(x,y)\to P_0} \frac{|f(x,y)-L_{P_0}(x,y)|}{||(x,y)-P_0||} = 0$

Geometrically, if a circle about P_0 is drawn and radius is collapsed, distance between f and approximating plane will become 0 much faster, so fraction becomes 0.

If there is a radius $\delta > 0$ on the disk of radius δ centered at P_0, $f_x(x,y)$ and $f_y(x,y)$ are continuous at every point on the disk, then $f(x,y)$ is continuous at every point on the disk. Differentiability can not be established by the existence of partial derivatives at a point, it is sufficient to demonstrate continuous partials at a point.

The criteria for differentiability are that f_x, f_y exist and f is locally linear at P_0. The tangent plane exists here. However, it is sufficient to demonstrate differentiability by showing that both f_x and f_y are continuous on an open disk D to conclude that $f(x,y)$ is differentiable on D. Alternatively, the above limit definition can be used.

If a function $f:\mathbb{R}^m \to \mathbb{R}^n$ exists, then it is made up of n component functions on $\mathbb{R}^m$. The Jacobian matrix is as follows:

$$Df(x_1,\ldots x_m) = \begin{bmatrix}\nabla f_1(x_1,\ldots x_m)\\ \vdots \\ \nabla f_n(x_1,\ldots x_m)\end{bmatrix}$$

The number of rows n depends on components (codomain) of the output space whereas the domain determines columns (m). Thus, linear approximation to a function $f:\mathbb{R}^m \to \mathbb{R}^n$ at P_0 is

$$L_{P_0}(\mathbf{x}) = f(P_0) + \left(\text{matrix of partial derivatives at} P_0\right)(\mathbf{x}-P_0)$$

In parametric equations (paths), let $\mathbf{c}(t) = \left(x(t), y(t), z(t)\right)$ have continuous partials. Thus, the matrix of partials $D\mathbf{c}(t) = \begin{bmatrix}x'(t)\\ y'(t)\\ z'(t)\end{bmatrix}$. Is the vertical velocity vector. If the approximation $L_{t_0}(t)$ is found, it yields:

$$
\begin{aligned}
L_{t_0}(t) &= \mathbf{c}(t_0) + \begin{pmatrix} x'(t_0) \\ y'(t_0) \\ z'(t_0) \end{pmatrix} (t - t_0) \\
&= \mathbf{c}(t_0) + (t - t_0)\mathbf{c}'(t_0)
\end{aligned}
$$

Also, note that matrix of partials for a scalar valued function $f : \mathbb{R}^n \to \mathbb{R}$ is simply the gradient; there is 1 row and n columns in the gradient vector.

19 Derivative Rules

General rules

If $f, g : \mathbb{R}^m \to \mathbb{R}^n$ are differentiable at $P_0 \in \mathbb{R}^m$ for $c \in \mathbb{R}$:

- $h(\mathbf{x}) = cf(\mathbf{x}) \implies \mathrm{D}h(P_0) = cDf(P_0)$
- $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}) \implies \mathrm{D}h(P_0) = \mathrm{D}f(P_0) + \mathrm{D}g(P_0)$ (domain = codomain; sum of two $n \times m$ matrices)

If $f, g : \mathbb{R}^n \to \mathbb{R}$ are differentiable at $P_0 \in \mathbb{R}^n$:

- $h(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x}) \implies \mathrm{D}h(P_0) = f(P_0)\mathrm{D}g(P_0) + g(P_0)\mathrm{D}f(P_0)$ (no commutativity as matrix times scalar DNE)

If $f, g : \mathbb{R}^n \to \mathbb{R}$ are both differentiable at $P_0 \in \mathbb{R}^n$ and $g(P_0) \neq 0$:

- $\mathrm{D}\left(\frac{f}{g}\right)(P_0) = \frac{g(P_0)\mathrm{D}f(P_0) - f(P_0)\mathrm{D}g(P_0)}{g(P_0)^2}$

19.1 Chain Rule

Chain rule

If $g : \mathbb{R}^m \to \mathbb{R}^p$ and $f : \mathbb{R}^p \to \mathbb{R}^n$ and g is differentiable at $P_0 \in \mathbb{R}^m$ and the same for f at $g(P_0)$:

- $\mathrm{D}\Big(f \circ g\Big)(P_0) = \mathrm{D}f\big(g(P_0)\big)\mathrm{D}g(P_0)$

Thus, derivative of composites is a matrix product. $\mathrm{D}f(g(P_0)) \to n \times p$ and $\mathrm{D}g(P_0) \to p \times m$ so $\mathrm{D}(f \circ g)(P_0) \to n \times m$.

If $\mathbf{c} : \mathbb{R} \to \mathbb{R}^n$ is a path differentiable at $t_0 \in \mathbb{R}$ and $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\mathbf{c}(t_0)$, then:

$$
\frac{d}{dt} f(\mathbf{c}(t_0)) = \mathrm{D}f(\mathbf{c}(t_0)) D\mathbf{c}(t_0) = [\nabla f(x_1, x_2, \dots x_n)] \begin{bmatrix} \vdots \\ \mathbf{c}'(t_0) \\ \vdots \end{bmatrix} = \nabla f(\mathbf{c}'(t_0)) \cdot \mathbf{c}'(t_0)
$$

Thus, derivative of scalar valued function and path is dot product of gradient and velocity.

20 Gradient

20.1 Directional Derivative

Directional derivative of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ at some $P_0 \in \mathbb{R}^n$ in the direction of $\mathbf{v} \in \mathbb{R}^n$ is given by

$$\boxed{f_{\mathbf{v}}(P_0) = \nabla f(P_0) \cdot \frac{\mathbf{v}}{\|\mathbf{v}\|}}$$

Maximizing the directional derivative involves the dot product. Let $f : \mathbb{R}^n \to \mathbb{R}$ and $P_0 \in \mathbb{R}^n$ where $\nabla f(P_0)$ is defined. Also, let $\mathbf{u}$ be a unit vector such that $f_{\mathbf{u}}(P_0)$ is maximized. The directional derivative is

$$f_{\mathbf{u}}(P_0) = \nabla f(P_0) \cdot \mathbf{u} = \|\nabla f(P_0)\| \cos\theta$$

Since $\cos\theta \in [-1, 1]$, the directional derivative is maximized at $\cos\theta = 1 \implies \mathbf{u} || \nabla f(P_0) ||$.

Gradient fact

The gradient of a function at P_0 points in the direction of steepest ascent.

For a more general case given $f : \mathbb{R}^n \to \mathbb{R}$ and $P_0 \in \mathbb{R}^n$ where $\nabla f_{\mathbf{v}}(P_0)$ is maximized and $\nabla f(P_0)$ exists:

$$\begin{aligned}
\nabla f_{\mathbf{v}}(P_0) &= \nabla f(P_0) \cdot \frac{\mathbf{v}}{||\mathbf{v}||} \\
&= \nabla f(P_0) \cdot \frac{\nabla f(P_0)}{||\nabla f(P_0)||} \\
&= \frac{\nabla f(P_0) \cdot \nabla f(P_0)}{||\nabla f(P_0)||} \\
&= ||\nabla f(P_0)||
\end{aligned}$$

Thus, the length of the gradient vector tells how steep the ascent is in the direction of the steepest ascent.

20.2 Gradient Fields

A image in the codomain $\mathbb{R}^n$ where arrows originating at each point point in the steepest direction $\forall\, P \in \mathbb{R}^n$ where the partials exist. Are always perpendicular to level curves of f.

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $P \in \mathbb{R}^n$ a point where partials of f exist. Considering the level set $f(\mathbf{x}) = f(P)$, all values map to same codomain value as P. Also, let $\mathbf{c} : \mathbb{R}^n \to \mathbb{R}$ have an image entirely within this level set $f(\mathbf{x}) = f(P)$ and $\mathbf{c}(t_0) = P$. Then, the proof must show that $\nabla f(P) \cdot \mathbf{c}'(t_0) = 0$.

Note that since $h(t) = f(\mathbf{c}(t))$, $h'(t_0) = \nabla f\Big(\mathbf{c}(t_0)\Big) \cdot \mathbf{c}'(t_0) = \nabla f(P) \cdot \mathbf{c}'(t_0)$. Since the image of $\mathbf{c}(t)$ lies within $f(\mathbf{x}) = f(P)$, f is constant on the image of $\mathbf{c}$ (every point on $\mathbf{c}$ maps to $f(P)$ under f). Thus, $h(t) = f(\mathbf{c}(t)) \implies h'(t) = 0 \;\forall\; t$.

20.3 Sphere

Sphere is not a function as a point $P = (x, y)$ correspond to multiple z values. Can be viewed as a level surface in $\mathbb{R}^3$ for some $f : \mathbb{R}^3 \to \mathbb{R}$. The function is $f(x, y, z) = x^2 + y^2 + z^2$ and the level surface sphere is $x^2 + y^2 + z^2 = R^2$ where R is the radius. Given the point $(0,0,R)$, the gradient is upward along the z-axis. All curves passing through this point have a tangent vector perpendicular to $\nabla f(0,0,R)$ since gradients are perpendicular to level sets. It can be said that any vector perpendicular to $\nabla f(0,0,R)$ is a tangent to at least 1 curve in the sphere passing through $(0,0,R)$. Thus, the perpendicular space to $\nabla f(0,0,R)$ forms a plane of tangent vectors at $(0,0,R)$ and is a tangent plane. An expression for the tangent plane is

$$
\begin{aligned}
\nabla f(0,0,R) \cdot (x-0, y-0, z-R) &= 0 \\
f_x(0,0,R)x + f_y(0,0,R)y + f_z(0,0,R) &= 0 \\
2(0)x + 2(0)y + 2(R)(z-R) &= 0 \\
2Rz - 2R^2 &= 0 \\
z &= R
\end{aligned}
$$

Expression for the tangent plane to a level set of some $f : \mathbb{R}^3 \to \mathbb{R}$:

$$
\begin{aligned}
\nabla f(P_0) \cdot (x - x_0, y - y_0, z - z_0) &= 0 \\
f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0) + f_z(P_0)(z - z_0) &= 0 \\
f_x(P_0)x + f_y(P_0)y + f_z(P_0)z &= \nabla f(P_0) \cdot P_0
\end{aligned}
$$

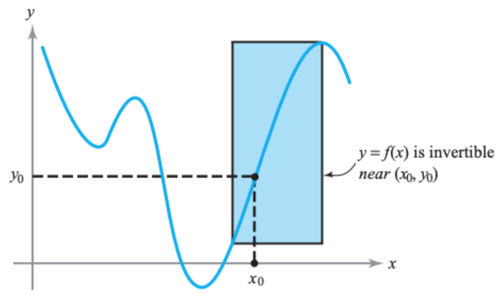
$$\boxed{f_x(P_0)x + f_y(P_0)y + f_z(P_0)z = \nabla f(P_0) \cdot P_0}$$

21 Implicit Function Theorem

21.1 Single variable

Is the explanation behind single-variable implicit differentiation. If $F : \mathbb{R}^2 \to \mathbb{R}$ is of class C^1 and there is a point $\mathbf{x}_0 = (a, b) \in \mathbb{R}^2$ such that $F(a, b) = 0$ and $F_y(a, b) \neq 0$, then it is guaranteed that

- There is some small $\delta > 0$ and some small $\epsilon > 0$ so that for $(x, y) \in \mathbb{R}^2$ with $|x - a| < \delta$ and $|y - b| < \epsilon$ there is a unique function $g : \mathbb{R} \to \mathbb{R}$ satisfying $F(x, g(x)) = 0$ and any point on the level set $F(x, y) = 0$ satisfying these 2 conditions on coordinates will have property $y = g(x)$. To summarize, near the point (a, b) points on $F(x, y) = 0$ lie on the graph of a unique function $y = g(x)$.
- g is of class C^1; differentiable so $g'(x)$ is continuous and $g'(x) = -\frac{F_x}{F_y}$



Special case

Given that $F : \mathbb{R}^{n+1} \to \mathbb{R}$ has continuous partials, let $(\mathbf{x}, z)$ denote the points in $\mathbb{R}^{n+1}$ where $\mathbf{x} \in \mathbb{R}^n$ and $z \in \mathbb{R}$. Assume that $(\mathbf{x}_0, z_0)$ satisfies

$$F(\mathbf{x}_0, z) = 0 \text{ and } \frac{\partial F}{\partial z}(\mathbf{x}_0, z) \neq 0$$

Then, $\mathbf{x}_0$ and z_0 are contained in a neighborhood such that there exists a unique $z = g(\mathbf{x})$ defined for $\mathbf{x}$ and z in these neighborhoods such that

$$F(\mathbf{x}, g(\mathbf{x})) = 0$$

It can also be said that $z = g(\mathbf{x})$ is continuously differentiable with derivative

$$\mathbf{D}g(\mathbf{x}) = -\frac{1}{\frac{\partial F}{\partial z}(\mathbf{x}, z)} \mathbf{D}_{\mathbf{x}} F(\mathbf{x}, z) \bigg|_{z=g(\mathbf{x})}$$

$\mathbf{D}_{\mathbf{x}} F$ denotes the partial of F with respect to $\mathbf{x}$ so that $\mathbf{D}_{\mathbf{x}} F = [\partial F/\partial x_1, \ldots, \partial F/\partial x_n]$. Thus,

$$\frac{\partial g}{\partial x_i} = -\frac{\partial F/\partial x_i}{\partial F/\partial z}, \quad i = 1, \ldots, n$$

21.2 Multivariable definition

If $F : \mathbb{R}^3 \to \mathbb{R}$ is of class C^1 and there exists $\mathbf{x}_0 = (a, b, c) \in \mathbb{R}^3$ such that $F(a, b, c) = 0$ and $F_z(a, b, c) \neq 0$ then is is guaranteed that

- There is some small $\delta > 0$ and $\epsilon > 0$ so that for $(x, y, z) \in \mathbb{R}^3$ with $\sqrt{(x-a)^2 + (y-b)^2} < \delta$ and $|z - c| < \epsilon$ there is a unique function $g : \mathbb{R}^2 \to \mathbb{R}$ that satisfies $F(x, y, g(x, y)) = 0$ and any point satisfying these inequalities on x, y, z-coordinates that lies on $F(x, y, z) = 0$ will have the property $z = g(x, y)$, so near the point (a, b, c) on the level set $F(x, y, z) = 0$ points can be seen as lying on $z = g(x, y)$.
- g is of class C^1 so it is differentiable and both partials exist and are continuous, and $g_x(x, y) = -\frac{F_x}{F_z}, g_y(x, y) = -\frac{F_y}{F_z}$.

The IFT can justify the existence of a tangent plane to a surface.

Let $F : \mathbb{R}^3 \to \mathbb{R}$ be of class C^1 and the surface $S = \{(x, y, z) \in \mathbb{R} \mid F(x, y, z) = c\}$ for some $c \in \mathbb{R}$ be the level set of F. The tangent plane is given by $F_x(P)x + F_y(P)y + F_z(P)z = \nabla F(P) \cdot \mathbf{P}$ for $P = (x_0, y_0, z_0)$. The plane is not defined when $\nabla F(P) = \mathbf{0}$.

Applying the IFT, this plane is tangent to S. If $\nabla F(P) \neq \mathbf{0}$ then at least one of $F_x, F_y, F_z \neq 0$ at P. Without loss of generality, let $F_z \neq 0$. Then, let $f(x,y,z) = F(x,y,z) - c$. Since $f_z(P) = F_z(P) \neq 0$, applying the IFT to f makes the conclusion that near P z is a differentiable function of x, y so that there is a unique function $g : \mathbb{R}^2 \to \mathbb{R}$ such that $z = g(x,y)$. The tangent plane to g at P is

$$\begin{aligned} z &= g(x_0, y_0) + g_x(x_0, y_0)(x - x_0) + g_y(x_0, y_0)(y - y_0) \\ &= z_0 + g_x(x_0, y_0)(x - x_0) + g_v(x_0, y_0)(y - y_0) \end{aligned}$$

It can then be shown that this plane is the same as given by the gradient approximation:

It follows that $f_x(P) = F_x(P)$ and $f_y(P) = F_y(P)$. Then,

$$\begin{aligned} z &= z_0 + g_x(x_0, y_0)(x - x_0) + g_y(x_0, y_0)(y - y_0) \\ z &= z_0 - \frac{F_x(P)}{F_z(P)}(x - x_0) - \frac{F_y(P)}{F_z(P)}(y - y_0) \\ F_z(P)z &= F_z(P)z_0 - F_x(P)x + F_x(P)x_0 - F_y(P)y + F_y y_0 \\ F_x(P)x + F_y(P)y + F_z(P)z &= F_z(P)z_0 + F_x(P)x_0 + F_y(P)y_0 \\ F_x(P)x + F_y(P)y + F_z(P)z &= \nabla F(P) \cdot P \end{aligned}$$

General implicit theorem

If $\det(A) \neq 0$ where A is the $m \times m$ matrix

$$\begin{bmatrix} \frac{\partial F_1}{\partial z_1} & \cdots & \frac{\partial F_1}{\partial z_m} \\ \vdots & \vdots & \\ \frac{\partial F_m}{\partial z_1} & \cdots & \frac{\partial F_m}{\partial z_m} \end{bmatrix}$$

then near the point $(\mathbf{x}_0, \mathbf{z}_0)$, the following system

$$\begin{aligned} F_1(x_1, \ldots, x_n, z_1, \ldots, z_m) &= 0 \\ F_2(x_1, \ldots, x_n, z_1, \ldots, z_m) &= 0 \\ &\vdots \\ F_m(x_1, \ldots, x_n, z_1, \ldots, z_m) &= 0. \end{aligned}$$

defines smooth and unique functions such that

$$z_i = k_i(x_1, \ldots, x_n) \quad (i = 1, \ldots, m)$$

where the derivatives are calculared with implicit differentiation.

21.3 IFT for vector-valued functions

Let $F:\mathbb{R}^{m+n}\to\mathbb{R}^n$ be of class C^1, then

$$F(x_1,x_2,x_3,...,x_m,z_1,z_2,...,z_n)=\Big(F_1(\text{input vector}),F_2(\text{input vector}),F_3(\text{input vector}),...,F_n(\text{input vector})\Big)$$

The inputs to F are the m-dimensional vector $\mathbf{x}_0=(x_1,x_2,\dots,x_m)$ and the n-dimensional vector $\mathbf{z}_0=(z_1,z_2,\dots,z_n)$. The conditions for IFT are met when

- $F(\mathbf{x}_0,\mathbf{z}_0)=\mathbf{0}$
- $\text{Det}\begin{pmatrix}\frac{\partial F_1}{\partial z_1} & \cdots & \frac{\partial F_1}{\partial z_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial F_n}{\partial z_1} & \cdots & \frac{\partial F_n}{\partial z_n}\end{pmatrix}\neq 0$

Then there is a unique C^1 function $g:\mathbb{R}^m\to\mathbb{R}^n$ such that for all points satisfying $F(\mathbf{v})=\mathbf{0}$ that are sufficiently near $P=(\mathbf{x}_0,\mathbf{z}_0)$ we can conclude $\mathbf{z}=g(\mathbf{x})$.

Inverse function theorem

This case is in the attempt to solve the system

$$\left.\begin{matrix}f_1(x_1,\dots,x_n)=y_1\\ \cdots\\ f_n(x_1,\dots,x_n)=y_n\end{matrix}\right\}$$

Solving this is the same as inverting the equations of this system. The condition is that in a neighborhood of some $\mathbf{x}_0$, $\det(A)\neq 0$ where A is the determinant of $\mathbf{D}f(\mathbf{x}_0)$ and $f=(f_1,\dots,f_n)$. $\det(A)$ (the Jacobian determinant) is given by

$$\left.\frac{\partial(f_1,\dots,f_n)}{\partial(x_1,\dots,x_n)}\right|_{\mathbf{x}=\mathbf{x}_0}=J(f)(\mathbf{x}_0)=\begin{vmatrix}\frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0)\\ \vdots & & \vdots\\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}_0)\end{vmatrix}$$

If this determinant is not 0, then the system can be solved for $\mathbf{x}=g(\mathbf{y})$ for $\mathbf{x}$ near $\mathbf{x}_0$ and $\mathbf{y}$ near $\mathbf{y}_0$ and g has continuous partials.

This is reasonable because a linear approximation acts in an invertible way locally, so both the derivative and function can be said to be invertible.

21.4 Example problem

Given a nonlinear system of equations:

$$xu+yvu^2=2$$
$$xu^3+y^2v^4=2$$

Implicit differentiation is done by treating u and v as functions of x, y. In this case, can find how a small change in x affects u when all variables are 1 (i.e. $(x, y, u, v) = (1,1,1,1)$)). First equation differentiation gives

$$\begin{aligned}
\frac{\partial}{\partial x}\left(xu + yvu^2\right) &= \frac{\partial}{\partial x}\left(2\right) \\
u + x\frac{\partial u}{\partial x} + yv\left(2u\frac{\partial u}{\partial x}\right) + yu^2\frac{\partial v}{\partial x} &= 0 \\
(x + 2uvy)\frac{\partial u}{\partial x} + yu^2\frac{\partial v}{\partial x} &= -u
\end{aligned}$$

Second equation gives

$$\begin{aligned}
\frac{\partial}{\partial x}\left(xu^3 + y^2v^4\right) &= \frac{\partial}{\partial x}(2) \\
x\left(3u^2\frac{\partial u}{\partial x}\right) + u^3 + y^2 4v^3\frac{\partial v}{\partial x} &= 0 \\
3xu^2\frac{\partial u}{\partial x} + y^2 4v^3\frac{\partial v}{\partial x} &= -u^3
\end{aligned}$$

Plugging in $(1,1,1,1)$ gives

$$\begin{aligned}
3u_x + v_x &= -1 \\
3u_x + 4v_x &= -1
\end{aligned}$$

Thus, $u_x = -\frac{1}{3}$. In this problem, F is defined as $F(x, y, u, v) = (xu + yvu^2 - 2, xu^3 + y^2v^4 - 2)$. Or,

$$\begin{aligned}
F_1(x, y, u, v) &= xu + yvu^2 - 2 \\
F_2(x, y, u, v) &= xu^3 + y^2v^4 - 2
\end{aligned}$$

The matrix to be checked for the condition is

$$\begin{bmatrix} \frac{\partial F_1}{\partial u} & \frac{\partial F_1}{\partial v} \\ \frac{\partial F_2}{\partial u} & \frac{\partial F_2}{\partial v} \end{bmatrix} = \begin{bmatrix} x + 2yvu & yu^2 \\ 3xu^2 & 4y^2v^3 \end{bmatrix} \implies \begin{bmatrix} 3 & 1 \\ 3 & 4 \end{bmatrix}$$

Since the determinant is 9, the IFT application is valid.