

MATH 119 - Calculus 2 for Engineering

Andy Zhang

Winter 2014

Contents

1	Approximation Methods	3
1.1	Linear Approximation	3
1.2	Bisection Method	3
1.3	Newton's Method(Not on exam)	3
1.4	Fixed-Point Iteration(Not on exam)	3
2	Polynomials	4
2.1	Polynomial Interpolation(Not on exam)	4
2.1.1	Problem	4
2.1.2	Steps	4
2.2	Taylor Polynomials	4
2.2.1	Idea	4
2.2.2	Remainder Theorem	5
2.2.3	Infinite Series	5
3	Convergence	6
3.1	Convergence Tests	6
3.1.1	Geometric Series	6
3.1.2	The Test for Divergence	6
3.1.3	The Integral Test	6
3.1.4	P-Series	6
3.1.5	The Comparison Test	6
3.1.6	The Limit Comparison Test	7
3.1.7	The Alternating Series Test	7
3.1.8	The Alternating Series Estimation Theorem	7
3.1.9	Absolute Convergence vs Conditional Convergence	7
3.1.10	Ratio Test	7
3.2	Radius of Convergence	7
3.3	Power Series Manipulation	8
3.4	Big-O Notation	8
3.4.1	Theorem	8
3.4.2	Evaluation of Limits using Taylor Polynomials and the Order Symbol	8
4	Multivariable Calculus	9
4.1	Multivariable Functions	9
4.2	Partial Derivatives	9
4.2.1	Clairaut's Theorem	9

4.3	Taylor Series: the Two-Variable Case	10
4.3.1	Tangent Plane and Differentials	10
4.4	Parametric Representation of Curves	10
4.5	Chain Rule	10
4.6	The Gradient Vector	10
4.7	Directional Derivatives	11
4.8	Even Nastier Chain Rule	11
5	Optimization Techniques	12
5.1	Unconstrained Optimization	12
5.2	Constrained Optimization	12
5.3	Identifying Absolute Extrema under Constraints	13
5.4	Integration of Scalar Fields	13
5.4.1	Evaluation of Double Integrals over Non-Rectangular Fields	13
5.4.2	Evaluation of Double Integrals in Polar Coordinates	14

Chapter 1

Approximation Methods

1.1 Linear Approximation

By using the slope at point (a, b) , we're estimating the value $f(x)$ where x is near a :

$$L(x) = y = f(a) + f'(a)(x - a)$$

1.2 Bisection Method

By taking advantage of the Intermediate Value Theorem, which states that if $f(a) < 0$ and $f(b) > 0$ and $f(x)$ is a continuous function, then there must exist a $c \in [a, b]$ such that $f(c) = 0$.

The idea is to use 2 points, bisect the interval into 2 intervals and check which one contains the root i.e. which one has a positive and a negative output.

1.3 Newton's Method(Not on exam)

If we can't solve $f(x) = 0$, then we can replace $f(x)$ with a linear approximation $L(x)$ and solve equation $L(x) = 0$ instead. This can be illustrated as an iterative formula which is described as:

The Linear Approximation: $L_{x_n}(x) = f(x_0) + f'(x_0)(x - x_0)$

The Iterative Formula: $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

1.4 Fixed-Point Iteration(Not on exam)

Theorem: Convergence of Fixed-Point Iteration Suppose that $f(x)$ is defined for all $x \in \mathbb{R}$, that it is differentiable everywhere, and that its derivative is always bounded(so that there are no points with vertical tangents). If the equation $f(x) = x$ has a solution(i.e. if $f(x)$ has a fixed point), and **if $|f'(x)| < 1$ for all values of x within some interval containing the fixed point**, then the sequence generated by letting $x_{n+1} = f(x_n)$ will converge, with *any* choice of x_0 .

Chapter 2

Polynomials

2.1 Polynomial Interpolation(Not on exam)

2.1.1 Problem

Suppose we are given $n + 1$ points and we would like to find a smooth curve which passes through all of them. The simplest such curve which passes through all of them of degree n .

2.1.2 Steps

1. Let's say there are $n + 1$ points. We define the polynomial as
$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$
2. Plug in each point into this polynomial and obtain n functions.
3. Incrementally find the difference between i and $i + 1$, where $\delta y_i = y_{i+1} - y_i$. As a result, this reduces the amount of variables in the polynomial
4. After isolating the first coefficient, it becomes simple to isolate each other coefficients using the derived equations.
5. We can then plug these coefficients into the original polynomial and obtain a result.

OR you can use a general formula:

$$y = y_0 + x\Delta y_0 + x(x-1)\frac{\delta^2 y_0}{2!} + \dots + x(x-1)(x-2)\dots(x-n+1)\frac{\Delta^n y_0}{n!}$$

Side Note: If each node is equidistant, then there's a simpler generalization:

$$y = y_0 + \frac{x-x_0}{h}\Delta y_0 + \frac{(x-x_0)(x-x_1)}{2!h^2}\Delta^2 y_0 + \dots + \frac{(x-x_0)(x-x_1)\dots(x-x_{n-1})}{n!h^n}\Delta^n y_0$$

2.2 Taylor Polynomials

2.2.1 Idea

The idea behind this is to use the secant line in the Linear Approximation and using Polynomial Interpolation to derive a polynomial that is similar to a function's shape due to its derivatives. By reaching the n^{th} derivative where n gets bigger and bigger, the polynomial resembles more and more to the function.

tl;dr: $P_{n,x_0}(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$

This becomes really handy when you need to do complicated stuff on functions such as integrals since polynomials are easy to integrate.

2.2.2 Remainder Theorem

Suppose that f has $n + 1$ derivatives at x_0 . Then

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + R_n(x)$$

where

$$R_n(x) = \int_{x_0}^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt$$

By applying some magic on this, we obtain Taylor's Inequality:

The error in using the n^{th} -order polynomial $P_{n,x_0}(x)$ as an approximation to $f(x)$ satisfies the inequality

$$|R_n(x)| \leq K \frac{|x-x_0|^{n+1}}{(n+1)!}$$

where $|f^{(n+1)}(t)| \leq K$ for all values of t between x_0 and x . This allows us to find the error margin of the Taylor Polynomial when estimating values.

2.2.3 Infinite Series

We can generalize some popular functions into infinite series due to a recurring pattern in their respective Taylor Polynomial:

- $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ (converges for $|x| < 1$)
- $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ (converges for all x)
- $\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$ (converges for all x)
- $\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$ (converges for all x)

Chapter 3

Convergence

3.1 Convergence Tests

3.1.1 Geometric Series

A geometric series has the form

$$\sum_{k=0}^{\infty} ar^k = a + ar + ar^2 + \dots$$

This is equivalent to

$$\sum_{k=0}^{\infty} ar^k = \lim_{n \rightarrow \infty} \frac{a(1-r^{n+1})}{1-r}$$

- If $|r| < 1$, the sequence r^n approaches zero, and so the above equality works
- If $|r| > 1$, the sequence r^n diverges, and so the series diverges as well.
- If $r = 1$, our series diverges since all of the $r^n = 1$ so the series becomes $a + a + a + \dots$
- If $r = -1$, then the series $r^n = (-1)^n$ diverges, and so does the series.

3.1.2 The Test for Divergence

If $\lim_{k \rightarrow \infty} a_k \neq 0$, then $\sum a_k$ diverges

3.1.3 The Integral Test

$\sum_{k=k_0}^{\infty} a_k$ converges iff $\int_{k_0}^{\infty} f(x)dx$ converges, where $f(k) = a_k$

3.1.4 P-Series

$\sum \frac{1}{k^p}$ converges if $p > 1$, and diverges if $p \leq 1$

3.1.5 The Comparison Test

Suppose we are given a series $\sum a_k$ (with all terms positive). If we can identify a second series $\sum b_k$ such that $a_k \leq b_k$ for all k , and $\sum b_k$ converges, then $\sum a_k$ also converges. Similarly, $a_k \geq b_k$, and $\sum b_k$ diverges, then $\sum a_k$ also diverges

3.1.6 The Limit Comparison Test

If $\lim_{k \rightarrow \infty} \frac{a_k}{b_k} = L$, where L is a non zero constant ($0 < L < \infty$), then $\sum a_k$ and $\sum b_k$ either both converge or diverge.

3.1.7 The Alternating Series Test

Consider the series $a_0 - a_1 + a_2 - \dots = \sum_{k=0}^{\infty} (-1)^k a_k$ where $a_k > 0$ for every k . If $\lim_{k \rightarrow \infty} a_k = 0$ and the sequence $\{a_k\}$ is decreasing, then the series converges

3.1.8 The Alternating Series Estimation Theorem

Consider a convergent alternating series $\sum (-1)^k a_k$. If we use the n^{th} partial sum s_n as an estimate of the sum, then the error satisfies the inequality $|s - s_n| \leq a_{n+1}$. That is, the truncation error is less than the first term omitted

3.1.9 Absolute Convergence vs Conditional Convergence

A series $\sum a_k$ is said to be absolutely convergent if the series $\sum |a_k|$ is convergent.

A series $\sum a_k$ is said to be conditionally convergent if it is convergent but the series $\sum |a_k|$ is divergent

3.1.10 Ratio Test

Suppose $\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = L$

- If $L < 1$, then the series $\sum a_k$ is absolutely convergent
- If $L > 1$, then the series $\sum a_k$ is divergent
- If $L = 1$, then the test fails, could be anything

3.2 Radius of Convergence

A series converges absolutely if

$$\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| = |x - x_0| \lim_{k \rightarrow \infty} \left| \frac{c_{k+1}}{c_k} \right| < 1$$

By letting $R = \lim_{k \rightarrow \infty} \left| \frac{c_k}{c_{k+1}} \right|$, we can establish that the series converges if $|x - x_0| < R$. We can state the following:

- If $R = 0$, then the series converges only at $x = x_0$
- If $R = \infty$ the the series converges for all x
- If $0 < R < \infty$, then the series converges absolutely for $x \in (x_0 - R, x_0 + R)$ and diverges for $x < x_0 - R$ and for $x > x_0 + R$. At the two points $x = x_0 \pm R$ gives no conclusion so we need to test them individually

We refer to R as the radius of convergence

3.3 Power Series Manipulation

If the series $\sum c_k(x - x_0)^k$ has radius of convergence R , then we can

- differentiate it (term by term)
- integrate it (term by term)
- multiply through by a constant (term by term)
- add it (term by term) to another series of radius of convergence $\geq R$

and the result will also have radius of convergence R . Note that to add, both series must have the same starting value for the index

3.4 Big-O Notation

3.4.1 Theorem

Given two functions f and g , we say that " f is of order g as $x \rightarrow x_0$ " and write

$$f(x) = \mathcal{O}(g(x)) \text{ as } x \rightarrow x_0$$

if there exists a constant $A(> 0)$ such that

$$|f(x)| \leq A|g(x)|$$

on some interval around x_0 (although the point x_0 itself may be excluded from the interval, since the idea is to describe the behaviour of f in the limit as we approach x_0)

This serves a purpose with Taylor's Inequality. Rather than calculating the estimated remainder, we can define an expression:

$$R_n(x) = \mathcal{O}((x - x_0)^{n+1}) \text{ as } x \rightarrow x_0$$

This serves a huge purpose when manipulating Taylor polynomials using addition, subtraction and especially multiplication where many expressions get absorbed by the Big O expression.

3.4.2 Evaluation of Limits using Taylor Polynomials and the Order Symbol

By replacing complex elements in a limit with Taylor polynomials with a Big-O expression as a remainder, it simplifies the calculation without affecting the answer.

Chapter 4

Multivariable Calculus

4.1 Multivariable Functions

Basically Calculus but with more than 2 variables. To visualize these functions, we will use level curves. We can set $z = 0, 1, 2, 3, ..$ and draw a graph for each value.

When checking for the origin, we can test the function with different values. Say we're checking for the origin. If $x = 0$, then we determine the limit as $y \rightarrow 0$ and vice versa. We also need to check if this limit corresponds to the same when we let $y = x$. We can also let $y = f(x)$ where $f(x)$ is an arbitrary function and let it approach the origin.

In order to prove that a limit exists, we need to prove that it exists along every conceivable route towards the point in question. The best tool would be Squeeze Theorem.

4.2 Partial Derivatives

The partial derivative of $f(x, y)$ with respect to x at the point (a, b) is

$$f_x(a, b) = \lim_{h \rightarrow 0} \frac{f(a+h, b) - f(a, b)}{h}$$

if this limit exists (otherwise f is not differentiable). Similarly, the partial derivative of $f(x, y)$ with respect to y at the point (a, b) is

$$f_y(a, b) = \lim_{h \rightarrow 0} \frac{f(a, b+h) - f(a, b)}{h}$$

providing that this limit exists

While differentiating, we essentially treat one of the variables as a constant

4.2.1 Clairaut's Theorem

Clairaut's Theorem states that if f_x, f_y and f_{xy} exist near (a, b) and if f_{xy} is continuous at (a, b) , then $f_{xy}(a, b)$ also exists, and in fact $f_{xy}(a, b) = f_{yx}(a, b)$

4.3 Taylor Series: the Two-Variable Case

Equation explains itself:

$$\begin{aligned} f(x, y) &= f(x_0, y_0) \\ &+ f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ &+ \frac{1}{2!} f_{xx}(x_0, y_0)(x - x_0)^2 + f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + \frac{1}{2!} f_{yy}(x_0, y_0)(y - y_0)^2 \\ &\dots \end{aligned}$$

4.3.1 Tangent Plane and Differentials

If all we need is a linear approximation, then we just need the first order polynomial:

$$f(x, y) \approx f(P_0) + f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0)$$

This can be summarized as:

$$df = f_x dx + f_y dy$$

4.4 Parametric Representation of Curves

A parametric equation of a curve is its representation through equations which express the coordinates of the points of the curve as functions of a variable called a parameter. For example:

$$\begin{aligned} x &= \cos t \\ y &= \sin t \end{aligned}$$

This can also be defined as

$$\begin{aligned} \vec{r}(t) &= \langle x(t), y(t) \rangle \\ \vec{r}'(t) &= \langle x'(t), y'(t) \rangle \end{aligned}$$

4.5 Chain Rule

In the case where f depends on more than 1 variable. As a result, we need to differentiate in terms of each of them as the following:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}$$

We cannot differentiate z in terms of t which is why we can split $\frac{dz}{dt}$ into $\frac{\partial z}{\partial x} \frac{dx}{dt}$ and the same for y

4.6 The Gradient Vector

We define the gradient vector as:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$$

Using this, we can redefine the chain rule as:

$$\begin{aligned} \frac{dz}{dt} &= \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \\ &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt} \right) \\ &= \nabla f(\vec{r}(t)) \cdot \vec{r}'(t) \end{aligned}$$

4.7 Directional Derivatives

The directional derivative of $f(x, y)$ in the direction of a unit vector $\vec{u} = (u_1, u_2)$ at the point $\vec{a} = (a, b)$ is denoted by $D_{\vec{u}}f(a, b)$ and defined as:

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

Since \vec{u} is a unit vector, this means that

$$\nabla f \cdot \vec{u} = \|\nabla f\| \cos \theta$$

This means that $D_{\vec{u}}f$ has maximum value when $\theta = 0$, meaning it is parallel with \vec{u}

4.8 Even Nastier Chain Rule

Let's say that f is a function of x and y . But what if x and y are in terms of s and t ? When differentiating in terms of s , we can just treat t as a constant and vice versa.

But what if we need the 2nd derivative? This gets pretty nasty though. We would essentially redifferentiate it and get something huge...(I'm not LaTeXing that..)

Essentially, keep differentiating until for each derivative you have to do, the function is in terms of whatever you're deriving it in terms of, i.e. if you're deriving f in terms of s but f is in terms of x and x is in terms of s , then derive f in terms of x multiplied by x derived in terms of s

Chapter 5

Optimization Techniques

5.1 Unconstrained Optimization

A function $f(x, y)$ has a local maximum at (x_0, y_0) if $f(x_0, y_0) \geq f(x, y)$ for all (x, y) in some disk centered at (x_0, y_0)

It's worth noting that a critical point is not necessarily an extremum unlike in single variable functions. If it's neither a maximum or minimum, we call this point a **saddle point**. We can check this point using the following test:

The Second Derivative Test for Local Extrema: Suppose P_0 is a critical point of an infinitely differentiable function $f(x, y)$. Let $D(x, y) = f_{xx}f_{yy} - (f_{xy})^2$

- If $D(P_0) > 0$, then f has an extremum at P_0 . If so, if $f_{xx}(P_0) < 0$, then this extremum is a maximum and vice versa
- If $D(P_0) < 0$, then this point does not have an extremum at P_0 (that is, it has a saddle point instead)
- If $D(P_0) = 0$, then the test gives no conclusion

5.2 Constrained Optimization

We follow these guidelines to find extremums:

- We know that ∇f is always orthogonal to the level of curves of f
- If we view the constraint curve as being one particular level curve of a function $g(x, y)$, then ∇g will be orthogonal to the constraint curve
- Therefore, any point where the constraint curve touches a level curve of f tangentially, it must be that ∇g is parallel to ∇f

This is the motivation of Lagrange:

To find the critical points of $f(x, y)$ subject to a constraint $g(x, y) = K$ (where K is a constant), find the values of x and y for which

$$\nabla f = \lambda \nabla g \text{ and } g(x, y) = K$$

Notes:

1. If the constraint is not given in the form $g(x, y) = K$, then we must rewrite it
2. The proportionality constant λ is called the *Lagrange Multiplier*. We normally don't care what the value of this is, unless we calculate it in order to solve for x and y . However, it does have an interesting interpretation; it can be shown that λ represents the rate of change of f with respect to the constrained value.

5.3 Identifying Absolute Extrema under Constraints

Lagrange's method only locates critical points without telling us whether they're local maxima, minima or inflection points.

In order to determine whether they're absolute maxima or minima, we'd follow these steps:

- If the constraint curve is closed and smooth, then one might think that we would simply have to compare the values of $f(x, y)$ at each of the points we've found by solving the equation in Lagrange's Method. The extrema could also occur at points where $\nabla g = \vec{0}$ and so we must determine if there are any such points, evaluate f at each one, and include those values in our list of values to compare
- If the curve has endpoints, then we must calculate the values of $f(x, y)$ at those points as well, and include them in the comparison
- If the curve is not closed, and has no endpoints, then $f(x, y)$ may not even be bounded. We'll need to consider the limits of $f(x, y)$ in the two directions along the curve to determine whether we have any absolute extrema at all

These methods do work for more than 2 variables as well. We'd let $g_1(x, y, z) = K_1$ and $g_2(x, y, z) = K_2$ and set $\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2$

5.4 Integration of Scalar Fields

It seems intuitive to see that since there's a partial derivative, there's partial integration. When integrating, we predefine the variable we're integrating the function in terms of. The only twist is that the constant of integration is in terms of the variable we set as constant.

When integrating over a scalar field, we partition the field in many tiny rectangles. Based on this, we define the double integral over a continuous function $f(x, y)$ over a rectangular region as

$$\int_R^{max} f(x, y) dA = \lim_{m, n \rightarrow 0} \sum_{i=1}^m \sum_{j=1}^n f(x_i^*, y_j^*) \Delta A$$

This can be simplified as:

$$\int \int f(x, y) dx dy$$

Integrating over rectangular fields should be straight forward

5.4.1 Evaluation of Double Integrals over Non-Rectangular Fields

We redefine the double integral as:

$$\int_R f(x, y) dA = \int_a^b \int_{g(x)}^{h(x)} f(x, y) dy dx$$

Since the limits of integrations are now functions, it's not possible to just switch the order of integration.

5.4.2 Evaluation of Double Integrals in Polar Coordinates

When the domain is circular, it is extremely useful to convert the coordinates to polar. An easy way to identify a situation would be to see an expression with $x^2 + y^2$ of some sort.

We would be integrating the function $f(r \cos \theta, r \sin \theta)$. However, we do need to change our differentials from $dx dy$ to $dr d\theta$. Through a bit of math, we see that:

$$\int \int_{R_{xy}} f(x, y) dx dy = \int \int_{R_{r\theta}} f(r \cos \theta, r \sin \theta) r dr d\theta$$

How did this happen? We use the **Change-Of-Variable** Formula, which is defined as the following:

Suppose that the variable x and y are related to the variables u and v by the equations $x = x(u, v)$ and $y = y(u, v)$. Then:

$$\int \int_{R_{xy}} f(x, y) dx dy = \int \int_{R_{uv}} f[x(u, v), y(u, v)] \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv$$

where

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$