

Math Camp Notes

Chris Ackerman

September 11, 2020

Abstract

These are my notes from UCLA summer math camp with Ichiro Obara. I've copied down the definitions that I thought would be useful throughout my PhD studies, usually because I learned something new, I think the definition will be used in coursework, or because it's something I was supposed to learn a long time ago but keep forgetting. If you have any questions, or if you notice any errors, please submit a pull request or email christopherackerman104@gmail.com.

Contents

1	Single Variable Calculus I	2
1.1	Rules for Taking Derivatives of Simple Analytic Functions	2
1.2	Exercises and Solutions	3
2	Single Variable Calculus II	6
2.1	Linear Approximation	6
2.2	Higher Order Derivatives	7
2.3	Exercises and Solutions	8
3	Introduction to Optimization	9
3.1	Constrained Optimization	10
3.2	Exercises and Solutions	11

1 Single Variable Calculus I

Definition (Graph). The set of (x, y) that f passes through is called the *graph* of f .

Definition (Convergent Sequence). A sequence x_{nn} *converges to* $x^* \in \mathbb{R}$ if, for any $\varepsilon > 0$, there exists an integer N such that $|x_n| \leq K$ for every n .

Definition (Continuous Function). Function $f : X \rightarrow \mathbb{R}$ is *continuous at* $x \in X$ if $x_n \rightarrow x$ for $x_n \in X \implies f(x_n) \rightarrow f(x)$. f is *continuous* if it is continuous at every x in its domain X .

Definition (Differentiable Function and Derivative). A function f on (a, b) is *differentiable* at $x \in (a, b)$ if $\frac{f(x+\Delta x)-f(x)}{\Delta x}$ converges to the same number for any sequence $\Delta x (\neq 0)$ such that $|\Delta x| \rightarrow 0$. This number is the *derivative* of f at x . f is differentiable on (a, b) if it is differentiable at every $x \in (a, b)$.

1.1 Rules for Taking Derivatives of Simple Analytic Functions

Definition (Product Rule).

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$$

Definition (Quotient Rule).

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$$

Definition (Chain Rule). Consider a composite function

$$h(x) = g(f(x)).$$

If f is differentiable at x and g is differentiable at $f(x)$, then h is differentiable at x and its derivative is given by

$$h'(x) = g'(f(x)) \cdot f'(x).$$

Definition (Inverse Function). The *inverse* function f^{-1} of f is defined for each x in the range of f as the unique value that satisfies

$$f(f^{-1}(x)) = x.$$

If f is differentiable and has an inverse f^{-1} , then f^{-1} is differentiable and the derivative of f^{-1} is the reciprocal of f' .

Definition (Elasticity). *Elasticity* measures the percentage change of a variable with respect to a percentage change of another variable. Suppose that two variables, x and y , satisfy $y = f(x)$. The *x-elasticity of y* at (x_0, y_0) is given by

$$\lim_{\Delta x \rightarrow 0} \frac{\frac{\Delta y}{y_0}}{\frac{\Delta x}{x_0}} = \lim_{\Delta x \rightarrow 0} \frac{\frac{f(x_0+\Delta x)-f(x_0)}{f(x_0)}}{\frac{\Delta x}{x_0}} = f'(x_0) \frac{x_0}{f(x_0)}.$$

We often express variables in natural logs to find the elasticity more easily. In logs, the *x-elasticity of y* is

$$\frac{d \ln y}{d \ln x}.$$

1.2 Exercises and Solutions

Question 1. Let $\{x_n\}_n$ and $\{y_n\}_n$ be two convergent sequences. Show that

1. $\lim(x_n y_n) = \lim x_n \lim y_n$,
2. $\lim\left(\frac{x_n}{y_n}\right) = \frac{\lim x_n}{\lim y_n}$, assuming $\lim y_n \neq 0$.

Answer.

1.

Note that

$$xy - x_n y_n = x(y - y_n) + (x - x_n)y_n.$$

So,

$$|xy - x_n y_n| \leq |x(y - y_n)| + |(x - x_n)y_n|. \quad (\text{Triangle Inequality})$$

Since $\{y_n\}_n$ is a convergent (thus bounded) sequence, there exists a K such that

$$\begin{aligned} |(x - x_n)y_n| &\leq |x - x_n||y_n| \\ &\leq |x - x_n|K. \end{aligned}$$

Do we need to repeat the argument for K here for $\{x_n\}$?

Then,

$$|x(y - y_n)| \leq |x||y - y_n| \rightarrow 0,$$

and

$$|(x - x_n)y_n| \leq |x - x_n|K \rightarrow 0.$$

Therefore,

$$|xy - x_n y_n| \rightarrow 0.$$

2.

Choose m so that $|y_n - y| < \frac{1}{2}|y|$ for $n \geq m$. Then,

$$|y_n| > \frac{1}{2}|y|. \quad (n \geq m)$$

Given $\varepsilon > 0$, there is an integer $N > m$ such that $n \geq N$ implies

$$|y_n - y| < \frac{1}{2}|s|^2\varepsilon.$$

Hence, for $n \geq N$,

$$\begin{aligned} \left| \frac{1}{y_n} - \frac{1}{y} \right| &= \left| \frac{y_n - y}{y_n y} \right| \\ &< \frac{2}{|y|^2} |y_n - y| \\ &< \varepsilon. \end{aligned}$$

Substitute this result into the previous proof to get the desired result.

□

Question 2. Show that a bounded increasing sequence $x_1 \leq x_2 \leq \dots \leq K < \infty$ must be a convergent sequence (use Bolzano-Weierstrasse).

Answer.

Since $\{x_n\}_n$ is bounded between x_1 and K , it has a convergent subsequence by BWT. Let $\{x_n\}_k$ denote this subsequence, and let x^* be its limit, and note that $x^* \leq K$. By the definition of a limit, for any $\varepsilon > 0$, there exists a \hat{k} such that

$$|x_{n(k)} - x^*| < \varepsilon \text{ for } k \geq \hat{k}.$$

Because the original sequence is increasing, this also holds for any $n \geq n(k)$ in the original sequence, completing the proof.

□

Question 3. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ be continuous functions. Show that $h(x) \equiv g(f(x))$ is a continuous function.

Answer. Start with $f(x)$ and take the limits one at a time. Use continuity at each step to show that the composite function h is also continuous.

□

Question 4. Derive the price elasticity of $q = -\frac{p}{3} + 8$ as a function of p . Note that the answer will only be defined for $p \in (0, 24)$.

Answer.

Start with the definition,

$$\begin{aligned}
 \varepsilon &= f'(p) \frac{p}{f(p)} \\
 f(p) &= \frac{-p}{3} + 8 \\
 f'(p) &= -\frac{1}{3} \\
 \varepsilon &= \frac{-1}{3} \cdot \frac{p}{\frac{-p}{3} + 8} \\
 &= -\frac{1}{3} \cdot \frac{3p}{-p + 24} \\
 &= \frac{-p}{24 - p}
 \end{aligned}$$

□

Question 5. Derive the x -elasticity of $f(x) = 3x^2$ and show that it does not depend on x . More generally, discuss why any function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with constant elasticity can be expressed as Ax^B with some $A > 0$ and $B \in \mathbb{R}$.

Answer.

Let's follow the same steps as the last question.

$$\begin{aligned}
 \varepsilon &= f'(x) \frac{x}{f(x)} \\
 f(x) &= 3x^2 \\
 f'(x) &= 6x \\
 \varepsilon &= \frac{6x \cdot x}{3x^2} \\
 &= 2.
 \end{aligned}$$

For the second part of the question, I'm going to prove the generalized form of the previous result.

$$\begin{aligned}
 f(x) &= Ax^B \\
 f'(x) &= ABx^{B-1} \\
 \varepsilon &= \frac{ABx^{B-1} \cdot x}{Ax^B} \\
 &= \frac{ABx^B}{Ax^B} \\
 &= B.
 \end{aligned}$$

This approach is different from Ichiro's solution and goes in the opposite direction. It doesn't have the same intuition from the log transform, but I think it's still sufficient.

□

Question 6. Show that the x -elasticity of $f(x)g(x)$ is the sum of the x -elasticity of $f(x)$ and the x -elasticity of $g(x)$.

Answer.

Take the log transform of the function. More generally, this is a nice way to get separable forms of multiplicative equations.

$$\begin{aligned}\frac{d \ln fg}{d \ln x} &= \frac{d \ln f + d \ln g}{d \ln x} \\ &= \frac{d \ln f}{d \ln x} + \frac{d \ln g}{d \ln x}.\end{aligned}$$

□

2 Single Variable Calculus II

2.1 Linear Approximation

We can use the derivative of f at x_0 to construct a linear approximation of f around x_0 :

$$L(x : x_0) = f(x_0) + f'(x_0)(x - x_0).$$

This approximation is a straight line that is tangent to f at x_0 . How good is this approximation? We can move x by Δx and evaluate the “error”.

Definition (Approximation Error). For an approximation $L(x_0 + \Delta x : x_0)$, the *approximation error* $R(\Delta x; x_0)$ is

$$R(\Delta x; x_0) = |f(x_0 + \Delta x) - L(x_0 + \Delta x : x_0)|.$$

In the case of linear approximation,

$$R(\Delta x; x_0) = |f(x_0 + \Delta x) - f(x_0) - f'(x_0)\Delta x|.$$

$\frac{R(\Delta x; x_0)}{\Delta x}$ converges to 0 as $\Delta x \rightarrow 0$, so the error goes to 0 faster than Δx goes to 0. $L(x : x_0)$ is the best linear approximation of f around x_0 .

Definition (Newton’s Method). *Newton’s method* is an example of an algorithm that uses linear approximations.

1. Guess any point x_0 .
2. Find a solution for the linear approximation $L(x : x_0 = 0)$ instead of $f(x) = 0$, which gives $x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$.
3. Repeat this procedure to obtain x_1, x_2, \dots by $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.
4. If x_n converges to some number x^* , then x^* must satisfy $f(x^*) = 0$.

Theorem (Mean Value Theorem). Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is differentiable and f' is continuous. Pick any point $x \in \mathbb{R}$ and $\Delta x > 0$. Then there exists $t \in (x, x + \Delta x)$ such that

$$f(x + \Delta x) - f(x) = f'(t)\Delta x.$$

This is a baby step towards Taylor's Theorem. For intuition, suppose that x increases to $x + \Delta x > x$ and $f(x)$ moves to $f(x + \Delta x)$. The rate of change is $\frac{f(x + \Delta x) - f(x)}{\Delta x}$. For differentiable f , this rate should be somewhere between the maximum derivative and the minimum derivative of f on $(x, x + \Delta x)$. The mean value theorem says that, when f' is continuous, this rate is the derivative of f at some point $t \in (x, x + \Delta x)$.

Let's reexamine the relative speed of $R(\Delta x; x_0) \rightarrow 0$ vs. $\Delta x \rightarrow 0$. By the mean value theorem,

$$R(\Delta x; x_0) = |(f'(t) - f'(x_0))\Delta x|$$

for some $t \in (x, x + \Delta x)$. So $\frac{R(\Delta x; x_0)}{\Delta x}$ converges to 0 at the same speed $f'(t)$ converges to $f'(x_0)$. If f' is differentiable, then

$$|f'(t) - f'(x_0)| \sim |f''(x_0)(t - x_0)| \leq |f''(x_0)\Delta x|.$$

Intuitively, $R(\Delta x; x_0)$ converges to 0 at the speed of $(\Delta x)^2$ in this case.

2.2 Higher Order Derivatives

A derivative is itself a function, so it is possible to consider the derivative of a derivative, etc.

Definition (Continuously differentiable). *A function is **continuously differentiable** if its derivative is continuous. A function is **k-times continuously differentiable** if it's k-times differentiable and its k^{th} derivative is continuous. The set of k-times continuously differentiable functions is denoted by \mathcal{C}^k .*

Polynomials, exponential functions, and long functions are all C^∞ . C^0 denotes a continuous function.

Definition (Taylor Polynomials). *For any $f \in \mathcal{C}^n$,*

$$L^k(x : x_0) = f(x_0) + \sum_{i=1}^k \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i$$

*is the **k^{th} -order Taylor polynomial** of f around x_0 , where $f^{(i)}$ is the i^{th} derivative of f .*

Higher order polynomials are more flexible than linear functions, so they provide a better local approximation. Taylor's Theorem is a generalization of the Mean Value Theorem.

Theorem (Taylor's Theorem). *Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is a \mathcal{C}^k function and $(k + 1)$ -times differentiable. Pick any point $x \in \mathbb{R}$ and $\Delta x > 0$. Then there exists $t \in (x, x + \Delta x)$ such that*

$$f(x + \Delta x) = L^k(x + \Delta x : x) + \frac{f^{(k+1)}(t)}{(k+1)!} (\Delta x)^{k+1}.$$

If f is a C^2 function, then

$$R(\Delta x; x_0) = \left| \frac{f''(t)}{2} (\Delta x)^2 \right|,$$

so $R(\Delta x; x_0)$ converges to 0 at the speed of $(\Delta x)^2$. More generally, if f is a C^{k+1} function, then the difference between $f(x_0 + \Delta x)$ and the k^{th} order Taylor polynomial

$$R_k(\Delta x; x_0) = |f(x_0 + \Delta x) - L^k(\Delta x; x_0)|$$

converges to 0 at the speed of $(\Delta x)^{k+1}$.

This formula has an important implication for local behavior. Whether f is strictly increasing or decreasing *locally* around x_0 is determined by the sign of $f'(x_0)$. If $f'(x_0) = 0$, then local behavior is determined by the sign of $f''(x_0)$, etc. Intuitively, the local effect of each derivative dominates the local effect of all higher order derivatives.

2.3 Exercises and Solutions

Question 7. Let f be a C^1 function and suppose that $f'(x) > 0$ at x . Show that there exists $\varepsilon > 0$ such that f is strictly increasing in $(x - \varepsilon, x + \varepsilon)$. Is the converse true?

Answer.

This is a fairly straightforward application of Taylor's theorem where $\Delta x = \pm\varepsilon$.

$$\begin{aligned} f(x - \varepsilon) &= f(x) + \underbrace{f'(t)(-\varepsilon)}_{(-)} \\ f(x + \varepsilon) &= f(x) + \underbrace{f'(t)(+\varepsilon)}_{(+)} \end{aligned}$$

Combining these inequalities satisfies the definition of an increasing function,

$$f(x - \varepsilon) < f(x) < f(x + \varepsilon).$$

□

Question 8. *L'Hôpital's Rule.* Suppose f and g are differentiable at f' , $f(x') = g(x') = 0$, and $g'(x') \neq 0$. Show that $\lim_{x \rightarrow x'} \frac{f(x)}{g(x)} = \frac{f'(x')}{g'(x')}$.

Answer. I'm going to use slightly different notation because I find x' cumbersome. This is a straightforward application of the definition of a first-order Taylor expansion.

$$\begin{aligned} f(x) &= f(x) + f'(x)\Delta x \\ g(x) &= g(x) + g'(x)\Delta x \\ \lim_{\Delta x \rightarrow 0} \frac{f(x)}{g(x)} &= \frac{\overbrace{f(x) + f'(x)\Delta x}^0}{\underbrace{g(x) + g'(x)\Delta x}_0} \\ &= \frac{f'(x)\Delta x}{g'(x)\Delta x} \\ &= \frac{f'(x)}{g'(x)} \end{aligned}$$

□

Question 9. For a C^2 function, verify that

$$L^2(x_0 + \Delta x : x_0) = f(x_0) + f'(x_0)\Delta x + \frac{f''(x_0)}{2}(\Delta x)^2$$

is the only 2nd order polynomial that satisfies

$$\frac{|f(x_0 + \Delta x) - L^2(x_0 + \Delta x : x_0)|}{(\Delta x)^2} \rightarrow 0$$

as $x \rightarrow 0$.

Answer. This solution is a proof by contradiction, in three parts. Since the denominator is going to zero, any value in the numerator that doesn't go to zero faster will blow up, and ruin convergence. Let $A + B\Delta x + C\Delta x^2$ be an arbitrary second order polynomial in Δx , and consider

$$\frac{|f(x + \Delta x) - (A + B\Delta x + C(\Delta x)^2)|}{(\Delta x)^2}.$$

1. First consider $A \neq f(x_0)$. If $A \neq f(x_0)$, then the numerator converges to $|f(x_0) - A| \neq 0$.
2. Now consider $B \neq f'(x_0)$. If $B \neq f'(x_0)$, then the numerator divided by Δx converges to $f'(x_0) - B \neq 0$.
3. Finally, consider $C \neq \frac{f''(x_0)}{2}$. In this case, the limit converges to $\left| \frac{f''(x_0)}{2} - C \right|$, which is only 0 if $C = \frac{f''(x_0)}{2}$.

Because of the rate at which the denominator goes to zero, the first two cases blow up, and the third case converges to the wrong number. \square

3 Introduction to Optimization

Throughout this section we assume that our functions are continuous (and their derivatives, etc. are continuous where necessary).

Definition (Simple Optimization Problem).

$$\max_{x \in [a, b]} f(x)$$

or

$$\max_{x \in \mathbb{R}} f(x) \text{ s.t. } x \geq a \wedge x \leq b$$

Example (Monopoly Profit Maximization). *Consider a monopoly firm with price setting power. If it sets price $p \geq 0$ for its product, it can sell $D(p)$ units. Each unit costs c to produce and there is a fixed cost F to start production. This firm's profit maximization problem (assuming they can shut down/choose not to produce) is*

$$\max \left\{ \max_{p \geq 0} (p - c)D(p) - F, 0 \right\}.$$

The profit maximizing price p^* satisfies the first order condition

$$D(p) + (p - c)D'(p) = 0.$$

Therefore the negative price elasticity of demand is the reciprocal of the profit margin at the optimal price p^* :

$$-\mathcal{E}_D(p^*) = \frac{p^*}{p^* - c}.$$

Theorem (First Order Condition). *If f achieves its maximum (or minimum) value at x^* on (a, b) , then $f'(x^*) = 0$.*

3.1 Constrained Optimization

When we optimize over a closed and bounded interval, we're guaranteed to get a maximum and a minimum point.

Theorem (Extreme Value Theorem). *If $f : \mathbb{R} \rightarrow \mathbb{R}$ is continuous, then it has a maximum point and a minimum point over any closed and bounded interval $[a, b]$.*

Definition (Necessary conditions for a maximum). *Suppose that x^* maximizes f on $[a, b]$. There are three possible cases, and each yields a set of necessary conditions that x^* must satisfy.*

1. $x^* \in (a, b) \implies f'(x^*) = 0$
2. $x^* = a \implies f'(a) \leq 0$
3. $x^* = b \implies f'(b) \geq 0$

Definition (Penalty/Shadow Price). *Consider the case with $x^* = b$. We can express the optimality condition with an equation*

$$f'(x) - \lambda = 0, \text{ where } \lambda = f'(b) \geq 0., \text{ where } \lambda = f'(b) \geq 0., \text{ where } \lambda = f'(b) \geq 0., \text{ where } \lambda = f'(b) \geq 0.$$

This has two useful economic interpretations.

1. λ is an artificial "penalty" or marginal cost that is associated with going beyond the constraint $x \leq b$. This interpretation allows us to turn the original constrained problem into an unconstrained problem with a penalty.
2. λ is the value of relaxing the constraint $x \leq b$. If b increases by $\Delta b > 0$, then the maximized value of f will increase by roughly $\lambda \Delta b$.

The first order condition is necessary for a maximum/minimum, but not sufficient. For continuous functions, though, we can look at the (sufficient) second order condition

Definition (Second Order Condition). *Suppose that f is C^2 and $f'(x^*) = 0$ at $x^* \in (a, b)$.*

1. If $f''(x^*) < 0$, then x^* is a strict local maximum for f .
2. If $f''(x^*) > 0$, then x^* is a strict local minimum for f .

Definition (Concavity and Convexity).

1. f is *concave* if

$$f((1 - \lambda)x + \lambda y) \geq (1 - \lambda)f(x) + \lambda f(y)$$

for every $x, y \in X$ and $\lambda \in [0, 1]$.

2. f is *convex* if

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y)$$

for every $x, y \in X$ and $\lambda \in [0, 1]$.

For *strictly concave* and *strictly convex*, make the inequalities strict.

Definition (Useful facts about concave functions).

If f is differentiable,

$$f \text{ is concave} \iff f(x) \leq f(x_0) + f'(x_0)(x - x_0) \quad \forall x \in X \text{ given any } x_0.$$

If f is twice differentiable,

$$f \text{ is concave} \iff f''(x) \leq 0 \quad \forall x \in X.$$

Similar results hold for convex functions; note that f is concave $\iff -f$ is convex.

Theorem (FOC Sufficiency). Suppose that f is a concave function on (a, b) and $f'(x^*) = 0$ at $x^* \in (a, b)$. Then f achieves its maximum value on (a, b) at x^* .

Proof. Since f is concave,

$$f(x) \leq f(x^*) + f'(x^*)(x - x^*) = f(x^*)$$

for every $x \in (a, b)$. □

Theorem (Another Sufficiency Condition). Suppose that $x^* \in (a, b)$ is the only point in (a, b) that satisfies the FOC $f'(x^*) = 0$. If $f''(x^*) < 0$, then x^* must be a maximum for f on (a, b) .

Does the strict inequality here mean that this is the unique maximum point for f on (a, b) ?

3.2 Exercises and Solutions

Question 10. Consider

$$\max_{x \in (a, b]} f(x).$$

What necessary condition must the optimal solution x^* satisfy?

Answer.

1. If $x^* \neq b$, $f'(x^*) = 0$.
2. If $x^* = b$, $f'(b) \geq 0$.

□

Question 11. Solve the following optimization problems.

1.

$$\max_{x \in \mathbb{R}_{++}} f(x) = \ln x - 3x$$

2.

$$\max_{x \in \mathbb{R}_+} -x^3 + 4x^2 - 5x - 2$$

Answer. 1.

$$\begin{aligned}f' &= \frac{1}{x} - 3 \\ \frac{1}{x} - 3 &= 0 \\ x &= \frac{1}{3}\end{aligned}$$

□

Question 12. A consumer spends x out of 100 to buy some product at price $p > 0$. Suppose that her utility when purchasing $x \in \left[0, \frac{100}{p}\right]$ units of the product is given by $u(x) = \frac{2x}{x+1} + 100 - px$. Find her utility-maximizing consumption $x(p)$ as a function of price.

Answer.

We know that she will spend her entire budget on the good if purchasing a single unit provides positive utility (local nonsatiation), or that she will not purchase any units if the price is too high (rationality). To figure out how much the consumer buys if she does purchase some of the good, take the derivative of demand and solve the FOC.

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{2 - p(x-1)^2}{(x+1)^2} \\ 0 &= \frac{2 - p(x-1)^2}{(x+1)^2}\end{aligned}$$

Since prices must be positive, take only the positive root. And because the consumer can choose not to buy anything, also include a maximization problem with the outside option in her demand function.

$$x(p) = \max \left\{ \frac{\sqrt{p^2 + 196p + 10404} - p + 102}{2p}, 0 \right\}$$

□

Question 13. Prove the following statements

1. If f is a concave function, then $-f$ is a convex function.
2. If f and g are concave, then $f + g$ is concave.
3. If f is concave and convex, then f must be a linear function.

Answer. 1. Take the definition, apply a negative sign, and the inequality flips (property of inequalities).

2. Take the definition, expand, and since we're adding two inequalities the inequality still holds in the same direction.

3.

The two inequalities together imply

$$\begin{aligned} f((1-\lambda)x + \lambda y) &= (1-\lambda)f(x) + \lambda f(y) \\ \frac{f(x + (y-x)) - f(x)}{\lambda(y-x)} &= \frac{f(y) - f(x)}{y-x} \end{aligned}$$

I need to review the rest of the proof more carefully since I don't see how the slope argument follows ...

□