Math Camp for Economists:

Optimization

Justin C. Wiltshire

Department of Economics
University of Victoria
Summer 2023

Function basics

In general, a function from a set X to a set Y is a rule which assigns to each element of X exactly one element of Y, where X is the "domain" of the function and Y is the "codomain" of the function

- We can denote this as $f : X \rightarrow Y$
  - $\rightarrow$ E.g. $f : X \rightarrow \mathbb{R}$ assigns a number in $\mathbb{R}$ (or $\mathbb{R}^1$) to each number in the set X

Consider a "one-variable" function, $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x - 2$

- $f(2) = 0$
- $f(1) = -1$

Consider another one-variable function, $g(x) = 2x^3$

- $g(2) = 16$
- $g(1) = 2$

Consider a "two-variable" function, $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $h(x, w) = x^2 - 2w$

- $h(2, 3) = -2$
- $h(3, 1) = 7$

Some important properties of functions

Some of the things we care about when considering a function:

(1) Whether it is increasing or decreasing

- A function f is increasing if $x_1 > x_2 \Rightarrow f(x_1) \geq f(x_2)$
- A function f is decreasing if $x_1 > x_2 \Rightarrow f(x_1) \leq f(x_2)$
- E.g. $f(x) = -x^3$ is a decreasing function

(2) The location of its local and global maxima and minima

- Where it changes from increasing to decreasing, and vice versa
- If $f(x)$ changes from decreasing to increasing at x_0, then the graph of f turns upward around the point $(x_0, f(x_0))$
  - $\rightarrow$ Then $(x_0, f(x_0))$ is a "local minimum" of f
  - $\rightarrow$ If $f(x) \geq f(x_0)$ for all x, then $(x_0, f(x_0))$ is a "global minimum" of f

(3) Where its curvature changes (its inflection points)

(4) Whether it's continuous

(5) Whether it's differentiable (and twice continuously differentiable)

Slope of a function

The "slope" of a function measures how it changes along some interval in its domain

- It describes direction and steepness (or the rate of change)

Consider a line in the Cartesian plane, $f(x) = mx + b$, where m and b are some constants. How does f change as x changes?

- If x changes from x_0 to x_1, then $\Delta x = x_1 - x_0$ is the change in x
- The associated change in f is $\Delta f(x) = f(x_1) - f(x_0)$: $m(x_1 - x_0)$
- The slope of f is then the change in f given the change in x: m $= \frac{\Delta f(x)}{\Delta x}$
  - $\rightarrow$ Note the slope of f is defined for any non-zero change in x
  - $\rightarrow$ Note the slope of f is constant because f is linear in x (it is a line!)

Slope of a non-linear function

Finding the slope of a linear function is intuitive and easy. For several reasons, this is not the case with non-linear functions:

- The slope of a non-linear function is not constant
- For any open interval in the domain of a function that is continuous and nonlinear on that interval, the slope of the function will vary along that interval
  - $\rightarrow$ Where along the interval should the slope of the function be measured?
- Maybe the function is discontinuous

Instead, use the slope of the tangent line to the graph of f at some point $(x_0, f(x_0))$

Slope of the tangent line to the graph of a function at a point

A tangent line to the graph of a function at a point x_0 'just barely touches' the graph at x_0

For some continuous nonlinear function $f(x)$:

- Draw a secant line joining $(x_0, f(x_0))$ and $(x_0 + h_1, f(x_0 + h_1))$, where h_1 is a small number
- Draw a secant line joining $(x_0, f(x_0))$ and $(x_0 + h_2, f(x_0 + h_2))$, where $h_2 < h_1$
  - $\rightarrow$ This is a better approximation to the desired tangent line than the first line
- Continue doing this choosing a sequence, $\{h_n\}$, which monotonically converges to zero
  - $\rightarrow$ A "monotonically increasing (decreasing) "sequence is bounded and non-decreasing (non-increasing)
- Then the slope is $\frac{f(x_0+h_n)-f(x_0)}{(x_0+h_n)-x_0} = \frac{f(x_0+h_n)-f(x_0)}{h_n}$

The derivative

- The "derivative" of f at x_0, written $f'(x_0)$ or $\frac{df}{dx}(x_0)$, is $f'(x_0) = \lim_{h \to 0} \dfrac{f(x_0 + h) - f(x_0)}{h}$ if the limit exists
  - $\rightarrow$ Can also be thought of as $\frac{df}{dx}(x_0)$ where $\frac{df}{dx} = \lim_{\Delta x \to 0} \dfrac{\Delta f(x)}{\Delta x}$
  - $\rightarrow$ When the limit exists, we say that f is "differentiable" at x_0 with derivative $f'(x_0)$

Some rules for calculating derivatives

Let a and b be real numbers, x be a real variable, and f, g, and h be differentiable functions at x_0:

- The derivative of $f(x) = a$ is $f'(x) = 0$ for all x
- The derivative of $f(x) = ax^b$ at x_0 is $f'(x_0) = abx_0^{b-1}$
- The derivative of $f(x) = e^x$ at x_0 is $f'(x_0) = e^{x_0}$
- The derivative of $f(x) = a^x$ at x_0 is $f'(x_0) = a^{x_0}\ln(a)$ if $a > 0$
- The derivative of $f(x) = \ln(x)$ at x_0 is $f'(x_0) = \frac{1}{x_0}$ if $x_0 > 0$
- $(af + bg)'(x_0) = ((af)' + (bg)')(x_0) = af'(x_0) + bg'(x_0)$, known as "the sum rule"
- $(af \cdot g)'(x_0) = a[f'(x_0)g(x_0) + f(x_0)g'(x_0)]$, known as "the product rule"
- $(\frac{f}{g})'(x_0) = \dfrac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$, known as "the quotient rule"
- If $f(x) = h(g(x))$, then $f'(x_0) = h'(g(x_0)) \cdot g'(x_0)$, known as "the chain rule"

A few definitions

A function $f : D \to \mathbb{R}$ is "continuous" at $x_0 \in D$ if, for any sequence $\{x_n\}$ which converges to x_0 in D, $f(x_n)$ converges to $f(x_0)$

- f is "continuous on a set" $U \in D$ if it is continuous at every $x \in U$
- f is continuous if it is continuous at every point in its domain

If $f(x)$ is a (continuous) differentiable function of x, then $f'(x)$ is another function of x. $f'(x)$ is itself continuous if the tangent line to the graph of f at $(x, f(x))$ changes continuously as x changes

- If $f'(x)$ is a continuous function of x, then the function f is "continuously differentiable", or C^1

Suppose f is C^1 at x_0. Then

- If $f'(x_0) > 0$, there is an open interval containing x_0 on which f is increasing
- If $f'(x_0) < 0$, there is an open interval containing x_0 on which f is decreasing
- The points at which $f'(x) = 0$ or f' is not defined are called "critical points" of f

Higher-order derivatives

Let f be a C^1 function on $\mathbb{R}$. Since $f'(x)$ is continuous on R, then f' may also have a derivative at x_0

- The derivative of $f'(x)$ at x_0 is the "second derivative" of f at x_0
- Written $f''(x_0)$ or $\dfrac{d}{dx}\dfrac{df}{dx}(x_0) = \dfrac{d^2 f}{dx^2}$

If f has a second derivative everywhere, and f'' is a continuous function of x, then f is "twice continuously differentiable", or C^2

Convexity of a function

When discussing the definiteness of a matrix we say a function f is convex if its domain is a convex set and, for all x, y in its domain and all $\lambda \in [0, 1]$, we have $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$

$\rightarrow$ We say f is "concave" if $-f$ is convex

We also said, earlier that $f'(x_0) > 0$ means there is an open interval containing x_0 on which f is *increasing*, and $f'(x_0) < 0$ means there is an open interval containing x_0 on which f is *decreasing*

Let f be a differentiable function:

- $f''(x) \geq 0$ on an interval I, such that f' is increasing on I, means f is convex on I
- $f''(x) \leq 0$ on an interval I, such that f' is decreasing on I, means f is concave on I

Exercises with functions

Exercise: Let f, g, and h be functions is convex if its domain is a convex set and, for all x, y in its domain and all $\lambda \in [0, 1]$, we have
$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$
$f(x) = x^2 + 2$, $g(x) = x$, and $h(x) = -x^3$

(1) For each function, pick 2–5 values of x (some positive, some negative) and find the value of the function at that value. Plot these (on different graphs), then approximately draw the function.

(2)

Lab Exercise 1

Exercise: Let $f(x) = ax^2 \ln(x)$, where a is a constant and x is a real variable

(1) Find $f'(x)$

(2) Find $f''(x)$

(3) Is f a C^2 function?

(4) Draw a graph of some function g that is C^1 on the open interval $(-3, 3)$, which at first increases, then decreases, then increases, then decreases. Indicate on the graph:

- Where is $g' > 0$?
- Where is $g' < 0$?
- Where is $g'' > 0$?
- Where is $g'' < 0$?
- Where is g increasing at an increasing rate?
- Where is g decreasing at an increasing rate?

Maxima and minima

We are often interested in the maxima ("max" points) and minima ("min" points) of functions

- Economists are often interested in maximizing profit or utility, or minimizing costs or $\sum_{i=1}^{n} \hat{u}_i^2$
- We defined minima and maxima earlier
- Minima and maxima can occur at the endpoint ("boundary") of the domain of f or in the "interior" of the domain of f
- If x_0 is a boundary max or min, then its derivative need not be zero
- If x_0 is an interior min or max of f, then
  - $\rightarrow$ x_0 is a "critical point" of f
  - $\rightarrow$ $f'(x_0) = 0$
- $f'(x_0) = 0$ is not sufficient to conclude x_0 is a max or min (or neither) of f
  (1) If $f'(x_0) = 0$ and $f''(x_0) < 0$, then x_0 is a max of f
  (2) If $f'(x_0) = 0$ and $f''(x_0) > 0$, then x_0 is a min of f
  (3) If $f'(x_0) = 0$ and $f''(x_0) = 0$, then x_0 could be a max or min of f, or neither

Max and min exercises

Exercise: Let $f(x) = x^4$, $g(x) = x^4 - 4x^3 + 4X^2 + 4$, and $h(x) = -4x^3$

(1) Find f', g', and h'. Are there any critical points for f, g, or h?

(2) Do you think there are max or min points for f, g, or h? Why or why not?

(3) Find f'', g'', and h''

(4) For each of f, g, and h, does this confirm your guess about max or min points? Why or why not?

(5) For any max or min points you found, do you think they are local or global extrema?

Global maxima and minima

In general, it's difficult to find a global max (min) of a function or prove any local max (min) is a global max (min)

Still, there are three cases when it's easier to find global extrema:

(1) When f has only one critical point in its domain;

(2) When $f'' > 0$ or $f'' < 0$ throughout the domain of f; or

(3) When the domain of f is a closed finite interval

Functions with only one critical point

Suppose that

a) the domain of f is an interval I (finite or infinite) in R,

b) x_0 is a local max of f, and

c) x_0 is the only critical point of f on I

Then x_0 is the global max of f on I

Functions with nowhere-zero second derivatives

If f is C^2 with an interval I as its domain, and if $f'' \neq 0$ on I, then f has at most one critical point in I that is a global max if $f'' < 0$ and a global min if $f'' > 0$

Functions whose domains are closed finite intervals

Let f be C^1 with a domain that is a closed and bounded interval $[a, b]$. We know an interior max or min of f must be a critical point of f. The only other candidates for a max or min are the two endpoints of the domain, $x = a$ and $x = b$

Then the global max of f can be found by doing the following:

(1) Compute the critical points of f by solving $f'(x) = 0$ for x in (a, b)

(2) Next, evaluate f at these critical points and at the endpoints, a and b, of its domain

(3) Finally, choose the point from among these that gives the largest value of f in step 2

Multivariable functions

When functions have several right-hand side variables it may be harder to picture their min and max

- Say we have a two-variable function $f(x, y)$. If we draw this, we need to evaluate f at each pair (x, y) and mark the point $(x, y, f(x, y))$ in $\mathbb{R}^3$

Exercise: Try to visually represent (draw) $f(x, y) = x^2 + y^2$

Level curves

An alternative and useful way of visualizing a function $f : \mathbb{R}^2 \to \mathbb{R}$ is to draw level curves (or "contours") in the plane

- For each (x, y), evaluate $f(x, y)$ to obtain a value, e.g. z_0
- Then sketch the locus in the $xy - plane$ of all other (x, y) pairs for which f also equals z_0
- This locus is called a "level curve" of f
- In economics we often consider production and utility functions with two variables. The associated level curves are isoquants and indifference curves, respectively

More generally, for $f : \mathbb{R}^n \to \mathbb{R}$, the set of n real variables for which f takes on a particular constant value c is the associated "level set"

Level curves exercise

Exercise:

(1) Draw the level curves of the function $f(x, y) = x^2 + y^2$ for $f = 1$, $f = 4$, and $f = 9$

(2) Imagine the area around UVic, which contains the two large hills of Mount Doug (PKOLS) and Mount Tolmie. Draw some level curves which you think approximate the elevation of the surrounding area in 20m intervals, where UVic is 60m above sea level (ASL), Mount Tolmie is 120m ASL and Mount Doug is 225m ASL

(3) Suppose you have a Cobb-Douglas production function $Q(L, K) = L^\alpha K^\beta$, with $\alpha = \beta = 1$. Draw the isoquants for $Q = 1$, $Q = 4$, and $Q = 8$

(4) Suppose you have a Cobb-Douglas utility function $U(x_1, x_2) = x_1^\alpha x_2^\beta$, with $\alpha = 1/2$ and $\beta = 1 - \alpha$. Draw the indifference curves for $U = 1$, $U = 2$, and $U = 4$

First-order derivatives of multivariable functions

With a one variable function, $f(x)$, we denoted the derivative as $\frac{df}{dx} = f'$. It is the instantaneous rate of change in f with respect to (wrt) x

When we have a multivariable function, $f : \mathbb{R}^n \to \mathbb{R}$, $f(x_1, ..., x_n)$, we often want to see how f changes when we change just one RHS variable at a time, holding the others constant

- The derivative of f wrt a single variable, x_i, is the "partial derivative" of f wrt x_i
- We denote it $\frac{\partial f}{\partial x_i}$ or f_i' or f_{x_i}' or $D_i f$
- This partial derivative of f wrt x_i at $\mathbf{x}^0 = \{x_1^0, ..., x_n^0\}$ in the domain of f is:

$$\frac{\partial f}{\partial x_i}(x_1^0, ..., x_n^0) = \lim_{h \to 0} \frac{f(x_1^0, ..., x_i^0 + h, ..., x_n^0) - f(x_1^0, ..., x_i^0, ..., x_n^0)}{h}$$

  if the limit exists

Exercise: Let $f(x_1, x_2) = 3x_1^4 x_2 + 2 \ln x_1 + 4x_2^2 + 5$. What is $\frac{\partial f}{\partial x_1}$ evaluated at $(2, 3)$? What about $\frac{\partial f}{\partial x_2}$?

Total differential

With a multivariable C^1 function, f, the effect of changing a single variable x_i ($dx_i = \Delta x_i$), holding all others constant, can be approximated by $\Delta f \approx \dfrac{\partial f(x_1^*, ..., x_n^*)}{\partial x_i} dx_i$

What if all the x_i change at once?

- The total change in f at $\mathbf{x}^* = (x_1^*, ..., x_n^*)$ is $\Delta f = f(x_1^* + \Delta x_1, ..., x_n^* + \Delta x_n) - f(x_1^*, ..., x_n^*)$
- We can approximate this by the "total differential", which is the sum of the individual one-variable changes:

$$\Delta f \approx df = \frac{\partial f(x_1^*, ..., x_n^*)}{\partial x_1} dx_1 + ... + \frac{\partial f(x_1^*, ..., x_n^*)}{\partial x_n} dx_n$$

- This can be written in matrix form as

$$df = \begin{bmatrix} \dfrac{\partial f(x_1^*, ..., x_n^*)}{\partial x_1} & \dots & \dfrac{\partial f(x_1^*, ..., x_n^*)}{\partial x_n} \end{bmatrix} \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix}$$

The gradient

This matrix of first derivatives of f evaluated at a point $\mathbf{x}^*$ is known as the "Jacobian" matrix of f at $\mathbf{x}^*$. Its transpose is called the "gradient":

$$Df(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}^*) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}^*) \end{bmatrix}$$

Think of it as a vector in $\mathbb{R}^n$ with tail at $\mathbf{x}^*$

At any point $\mathbf{x}$ in the domain of f at which $\nabla f(\mathbf{x}) \neq 0$, $\nabla f(\mathbf{x})$ points from $\mathbf{x}$ into the direction in which f increases most rapidly

Exercise with the gradient

Consider the production function $Q(K, L) = 4K^{3/4}L^{1/4}$

(1) Find the gradient of Q at $(10000, 625)$

(2) Interpret this

Higher-order derivatives of multivariable functions

We can also consider higher-order derivatives, as we do with one-variable functions

For a two-variable twice-differentiable function, $f(x, y)$, the second-order derivatives are:

- $f_{xx} = \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2}$
- $f_{xy} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial x \partial y}$
- $f_{yx} = \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \partial x}$
- $f_{yy} = \frac{\partial}{\partial y}\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial y^2}$

More generally, an n-variable twice-differentiable function, $g(x_1, ..., x_n)$, will have n^2 second derivatives of the form $\frac{\partial^2 g}{\partial x_i \partial x_j}$ (also written g_{ij} or g''_{ij}), where those for which $i \neq j$ are called the "cross partial derivatives"

The Hessian matrix

The second-order partial derivatives of an n-variable twice-differentiable function, $f(x_1, ..., x_n)$, can be arranged in an $n \times n$ "Hessian" matrix whose (i, j)th entry is $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}^*)$, written $D^2 f(\mathbf{x}^*)$ or $D^2 f_{\mathbf{x}}$:

$$D^2 f_{\mathbf{x}} = \begin{bmatrix} \dfrac{\partial f^2}{\partial x_1^2} & \dfrac{\partial f^2}{\partial x_2 \partial x_1} & \cdots & \dfrac{\partial f^2}{\partial x_n \partial x_1} \\ \dfrac{\partial f^2}{\partial x_1 \partial x_2} & \dfrac{\partial f^2}{\partial x_2^2} & \cdots & \dfrac{\partial f^2}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f^2}{\partial x_1 \partial x_n} & \dfrac{\partial f^2}{\partial x_2 \partial x_n} & \cdots & \dfrac{\partial f^2}{\partial x_n^2} \end{bmatrix}$$

If all n^2 of these exist and are continuous functions of $(x_1, ..., x_n)$, then f is C^2

Young's Theorem

If an n-variable function f is C^2 on an open region J in $\mathbb{R}^n$, then for all $\mathbf{x}$ in J and for each pair of indices i, j,

$$\frac{\partial f^2}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial f^2}{\partial x_j \partial x_i}(\mathbf{x})$$

Exercise: Let $Q(K, L) = AK^\alpha L^\beta$, where A is some constant, and K and L are real variables

(1) Find the gradient of Q

(2) Find the Hessian matrix

(3) Does Young's Theorem appear to hold?

Extrema of multivariable functions

Earlier we defined maxima and minima for a one-variable function. The general idea is the same for an n-variable function, $F\colon U \to \mathbb{R}$, whose domain U is a subset of $\mathbb{R}^n$:

(1) A point $\mathbf{X}^* \in U$ is a "global max" of F on U if $F(\mathbf{x}^*) \geq F(\mathbf{x})$ for all $\mathbf{x}$ in U, and a "strict global max" of F on U if $F(\mathbf{x}^*) > F(\mathbf{x})$ for all $\mathbf{x} \neq \mathbf{x}^*$ in U

(2) $\mathbf{x}^* \in U$ is a "local max" of F if there is a ball $B_r(\mathbf{x}^*)$ about $\mathbf{x}^*$ such that $F(\mathbf{x}^*) \geq F(\mathbf{x})$ for all $\mathbf{x} \in B_r(\mathbf{x}^*) \cap U$, and a "strict local max" of F if there is a ball $B_r(\mathbf{x}^*)$ about $\mathbf{x}^*$ such that $F(\mathbf{x}^*) > F(\mathbf{x})$ for all $\mathbf{x} \neq \mathbf{x}^*$ in $B_r(\mathbf{x}^*) \cap U$

(3) Reversing these inequalities leads to the definitions for global and local minima and strict global and local minima

Certain conditions on the first- and second-order derivatives are sufficient for determining local extrema

First-order conditions (FOCs) for an interior point to be a local extremum

Let $F: U \to \mathbb{R}$ be a C^1 function defined on a subset U of $\mathbb{R}^n$

- If $\mathbf{x}^*$ is an interior point of U, then necessary conditions for $\mathbf{x}^*$ to be a local max or min of F in U are that $\dfrac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0$ for $i = 1, ..., n$

Second-order sufficient conditions

Sufficient SOCs for an interior point to be a local extremum

If, in addition to the necessary FOCs, F is a C^2 function whose domain is an open set U in $\mathbb{R}^n$, then:

(1) If the Hessian $D^2F(\mathbf{x}^*)$ is a negative definite symmetric matrix, then $\mathbf{x}^*$ is a strict local max of F

- $\rightarrow$ $D^2F(\mathbf{x}^*)$ is "negative definite" if the n LPMS of $D^2F(\mathbf{x}^*)$ alternate in sign beginning with $|F''_{x_1x_1}| < 0$
- $\rightarrow$ Alternatively, $D^2F(\mathbf{x}^*)$ is "negative definite" if all n LPMs of $-D^2F(\mathbf{x}^*)$ are strictly positive

(2) If the Hessian $D^2F(\mathbf{x}^*)$ is a positive definite symmetric matrix, then $\mathbf{x}^*$ is a strict local min of F

- $\rightarrow$ $D^2F(\mathbf{x}^*)$ is "positive definite" if all n LPMs are strictly positive

(3) If the Hessian $D^2F(\mathbf{x}^*)$ is indefinite, then $\mathbf{x}^*$ is a "saddle point" and is neither a local min nor a local max of F

Second-order necessary conditions

Necessary SOCs for an interior point to be a local extremum

If, in addition to the necessary FOCs, F is a C^2 function whose domain U is in $\mathbb{R}^n$, then:

(1) If $\mathbf{x}^*$ is a local min of F, all the principal minors of the Hessian $D^2F(\mathbf{x}^*)$ are ≥ 0

(2) If $\mathbf{x}^*$ is a local max of F, then all the principal minors of the Hessian $D^2F(\mathbf{x}^*)$ of odd order are ≤ 0 and all those of even order are ≥ 0

Conditions for global extrema

Let $F\colon U \to \mathbb{R}$ be a C^2 function whose domain U is a convex open subset of $\mathbb{R}^n$

(1) The following three conditions are equivalent:

  (i) F is a concave function on U

  (ii) $F(\mathbf{y}) - F(\mathbf{x}) \leq DF(\mathbf{x})(\mathbf{y} - \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in U$

  (iii) $DF(\mathbf{x})$ is negative semidefinite for all $\mathbf{x} \in U$

(2) The following three conditions are equivalent:

  (i) F is a convex function on U

  (ii) $F(\mathbf{y}) - F(\mathbf{x}) \geq DF(\mathbf{x})(\mathbf{y} - \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in U$

  (iii) $DF(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in U$

(3) If F is a concave function on U and $DF(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x} \in U$, then $\mathbf{x}^*$ is a global max of F on U

(4) If F is a convex function on U and $DF(\mathbf{x}^*) = \mathbf{0}$ for some $\mathbf{x} \in U$, then $\mathbf{x}^*$ is a global min of F on U

Constrained maximization (general form)

We'll often want to find the optimal allocation of scarce resources (e.g. subject to constraints)

That is, we will often face an optimization problem of the form:

$$\max_{x1,\dots,x_n} f(x_1, \dots, x_n)$$

$$\text{subject to}$$

$$g_1(x_1, \dots, x_n) \leq b_1, \dots, g_k(x_1, \dots., x_n) \leq b_k,$$
$$h_1(x_1, \dots, x_n) = c_1, \dots, h_m(x_1, \dots, x_n) = c_m$$

where

- $f : x \subset X \in \mathbb{R}^n \to \mathbb{R}$ is the "objective function"
- $g_1, \dots, g_k$ are "inequality constraints" (usually "non-negativity constraints: $x_1 \geq 0, \dots, x_n \geq 0$)
  - $\to$ Technically these would be $-x_1 \leq 0, \dots, -x_n \leq 0$, but this is unnatural so don't for now
- $h_1, \dots, h_m$ are equality constraints

Example: Utility maximization (for a consumer)

- x_i is the amount of good i
- $f(x_1, ..., x_n)$ is usually written $U(x_1, ..., x_n)$
  - $\rightarrow$ Measures the consumer's utility/satisfaction with a consumption bundle $(x_1, ..., x_n)$
- p_i is the price of good i
- M (or I; or Y) is the consumer's income

The consumer's objective is to

$$\max_{x1,\ldots,x_n} U(x_1, ..., x_n)$$

subject to

$$p_1 x_1 + ... + p_n x_n \leq M,$$

$$x_1 \geq 0, ..., x_n \geq 0$$

where $p_1 x_1 + ... + p_n x_n \leq M$ is called the "budget constraint" and the non-negativity constrains mean consumers can't choose negative amounts of goods

Exercise: Equality constraints

Consider the basic utility maximization problem with $n = 2$. For now, assume the consumer spends all income with certainty, meaning $p_1x_1 + p_2x_2 = M$

Exercise: In the x_1x_2-plane, the non-negativity constraints restrict the constraint set, C, to $\mathbb{R}^2+$. The constraint set contains the set of bundles (x_1, x_2) which, given (p_1, p_2) and by the equality of the budget constraint, cost exactly M and which must contain the optimal bundle(s)

(1) Draw the the constraint set C in the x_1x_2-plane

(2) At $\mathbf{x_0}$, and keeping income constant, at what rate must this person trade x_2 to get another unit of x_1? (hint: set the total differential of the budget constraint to $dM = 0$)

(3) Draw some representative level curves of U if U is Cobb-Douglas with $\alpha, \beta \in (0, 1)$, $\alpha + \beta \leq 1$

- Note that with these preferences we always have an interior solution
- The highest level curve of U to touch C must be tangent to C at the constrained max $\mathbf{x}^*$
- This means the slope of the level set of U equals the slope of the constraint curve C at $\mathbf{x}^*$

(4) At $\mathbf{x_0}$, at what rate will this consumer be willing to trade x_2 for another unit of x_1 and be indifferent about it? This is the "marginal rate of substitution" (MRS) between x_1 and x_2 at $\mathbf{x_0}$

(5) Equate dx_2/dx_1 from (2) to that from (4).

More generally (but still with $n = 2$ and one equality constraint)

With an objective function $f(x_1, x_2)$ and one equality constraint $h(x_1, x_2) = c$:

- Set the total differential of $f(x_1, x_2)$ at $\mathbf{x}^*$ equal to zero, and solve for $\frac{dx_2}{dx_1}$
- Set the total differential of $h(x_1, x_2)$ at $\mathbf{x}^*$ equal to zero, and solve for $\frac{dx_2}{dx_1}$
- Equate the $\dfrac{dx_2}{dx_1}$
- Re write the result as $\dfrac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)} = \dfrac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)}$, and set $\dfrac{\frac{\partial f}{\partial x_1}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_1}(\mathbf{x}^*)} = \dfrac{\frac{\partial f}{\partial x_2}(\mathbf{x}^*)}{\frac{\partial h}{\partial x_2}(\mathbf{x}^*)} = \mu$
- Rewrite these equations as more general system:

$$\frac{\partial f}{\partial x_1}(\mathbf{x}) - \mu \frac{\partial h}{\partial x_1}(\mathbf{x}) = 0$$
$$\frac{\partial f}{\partial x_2}(\mathbf{x}) - \mu \frac{\partial h}{\partial x_2}(\mathbf{x}) = 0$$

Plus add the constraint:

$$h(x_1, x_2) - c = 0$$

The Lagrangian function

This system of three equations in three unknowns $(x1, x2, \mu)$ can be solved by finding the critical points of the "Lagrangian" function, L, where μ is a "Lagrange multiplier"

$$L(x_1, x_2, \mu) \equiv f(x_1, x_2) - \mu(h(x_1, x_2) - c)$$

This transformation effectively reduces the constrained optimization problem in two variables to an unconstrained optimization problem in three variables

Note: we have a "constraint qualification" on the constraint set, in that we need at least one of $dh/dx_1 \neq 0$ or $dh/dx_2 \neq 0$ at the constrained maximizer $\mathbf{x}^*$

Even more general: with n variables and m equality constraints

With n variables and m equality constraints, $h_1(x_1, \ldots, x_n) = c_1, \ldots, h_m(x_1, \ldots, x_n) = c_m$, the generalization of the constraint qualification is the Jacobian of the vector of equality constraints,

$$D\mathbf{h}(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial h_1}{\partial x_n}(\mathbf{x}^*) \\ \frac{\partial h_2}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial h_2}{\partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial h_m}{\partial x_n}(\mathbf{x}^*) \end{pmatrix}$$

$\mathbf{x}^*$ is called a critical point of $\mathbf{h} = (h_1, \ldots h_m)$ if $rank(D\mathbf{h}(\mathbf{x}^*)) < m$, <u>which we **don't** want</u>

$(h_1, \ldots, h_m)$ satisfies the "nondegenerate constraint qualification (NDCQ) at $\mathbf{x}^*$) if $rank(D\mathbf{h}(\mathbf{x}^*)) = m$

The Lagrangian with n variables and m equality constraints

Let $f(\mathbf{x}, h_1(\mathbf{x}), ..., h_m(\mathbf{x})$ be C^1 functions. Suppose we want to maximize $f(\mathbf{x})$ on the constraint set $C_h \equiv \{\mathbf{x} = (x_1, ..., x_n) \colon h_1(\mathbf{x}) = c_1, ..., h_m(\mathbf{x}) = c_m\}$

Suppose $\mathbf{x}^* \in C_h$ is a local max or min of f on C_h, and that $\mathbf{x}^*$ satisfies the NDCQ. Then $\exists\ \mu_1^*, ..., \mu_m^*$ such that $(\mathbf{x}^*, \mu^*)$ is a critical point of the Lagrangian

$$L(\mathbf{x}, \mu) = f(\mathbf{x}) + \mu_1[c_1 - h_1(\mathbf{x})] + ... + \mu_m[c_m - h_m(\mathbf{x})]$$

That is, we require:

$$\frac{\partial L}{\partial x_1}(\mathbf{x}^*, \mu) = 0, ..., \frac{\partial L}{\partial x_n}(\mathbf{x}^*, \mu) = 0$$
$$\frac{\partial L}{\partial \mu_1}(\mathbf{x}^*, \mu) = 0, ..., \frac{\partial L}{\partial \mu_m}(\mathbf{x}^*, \mu) = 0$$

… and with k inequality constraints

With n variables, m inequality constraints, and k inequality constraints $g_1(\mathbf{x}) \leq b_1, \ldots, g_k(\mathbf{x}) \leq b_k$, things are as they are with just m equality constraints except, in addition:

(1) The constraint set is
$C_{g,h} \equiv \{\mathbf{x} = (x_1, \ldots, x_n) \colon g_1(\mathbf{x}) \leq b_1, \ldots, g_k(\mathbf{x}) \leq b_k, h_1(\mathbf{x}) = c_1, \ldots, h_m(\mathbf{x}) = c_m\}$

(2) If the maximizer $\mathbf{x}^*$ is in the interior of $C_{g,h}$, then at least one $g_j(\mathbf{x}^*) < b_j$ for $j = 1, \ldots, k$ is "not binding" (alternatively, is "loose") at $\mathbf{x}^*$

(3) The Lagrange multipliers on the inequality constraints must be non-negative, $\lambda_1 \geq 0, \ldots, \lambda_k \geq 0$

(4) Either $\lambda_j \geq 0$ or $g_j(\mathbf{x}) \leq b_j$ must be binding $\Rightarrow \lambda_j[g_j(\mathbf{x}) - b_j] = 0$ for all $j = 1, \ldots, k$

   $\rightarrow$ This is called the "complementary slackness condition"

(5) Given $e \leq k$ binding inequality constraints and m equality constraints, the rank of the Jacobian of binding constraints is $e + m$

Most general: n variables, m equality constraints, k inequality constraints

Conditions on the Lagrangian with n variables, m equality constraints, and k inequality constraints provided the NDCQ is satisfied (that is the Jacobian of m equality and e binding inequality constraints is $e + m$)

We require:

a) $\dfrac{\partial L}{\partial x_1}(\mathbf{x}^*, \mu) = 0, ..., \dfrac{\partial L}{\partial x_n}(\mathbf{x}^*, \mu) = 0$

b) $\dfrac{\partial L}{\partial \mu_1}(\mathbf{x}^*, \mu) = 0, ..., \dfrac{\partial L}{\partial \mu_m}(\mathbf{x}^*, \mu) = 0$

c) $\lambda_1^* \geq 0, ..., \lambda_k^* \geq 0$

d) $g_1(\mathbf{x}^*) \leq b_1, ..., g_k(\mathbf{x}^*) \leq b_k$

e) $\lambda_1[g_1(\mathbf{x}^*) - b_1] = 0, ..., \lambda_k[g_k(\mathbf{x}^*) - b_k] = 0$

Exercise: maximization with equality and inequality constraints

Exercise: Suppose we want to maximize $f(x, y) = x - y^2$ subject to $x^2 + y^2 = 4, x \geq 0, y \geq 0$

(1) Check the NDCQ. At what $\mathbf{x}$ is the gradient zero? Is this point in the constraint set?

(2) Form the Lagrangian

(3) Find the <u>nine</u> FOCs

(4) Find the critical points of f

The bordered Hessian

We want to check whether the critical points of the Lagrangian are are minima or maxima (if either)

Construct the "bordered Hessian" to check the SOCs for constrained optimization with $e \leq k$ binding inequality constraints and m equality constraints:

$$H = \left(\begin{array}{cccccc|ccc}
0 & \dots & 0 & 0 & \dots & 0 & -\frac{\partial g_1}{\partial x_1} & \dots & -\frac{\partial g_1}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 & -\frac{\partial g_e}{\partial x_1} & \dots & -\frac{\partial g_e}{\partial x_n} \\
0 & \dots & 0 & 0 & \dots & 0 & -\frac{\partial h_1}{\partial x_1} & \dots & -\frac{\partial h_1}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & 0 & 0 & \dots & 0 & -\frac{\partial h_m}{\partial x_1} & \dots & -\frac{\partial h_m}{\partial x_n} \\
\hline
-\frac{\partial g_1}{\partial x_1} & \dots & -\frac{\partial g_e}{\partial x_1} & -\frac{\partial h_1}{\partial x_1} & \dots & -\frac{\partial h_m}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \dots & \frac{\partial^2 L}{\partial x_n \partial x_1} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
-\frac{\partial g_1}{\partial x_n} & \dots & -\frac{\partial g_e}{\partial x_n} & -\frac{\partial h_1}{\partial x_n} & \dots & -\frac{\partial h_m}{\partial x_n} & \frac{\partial^2 L}{\partial x_1 \partial x_n} & \dots & \frac{\partial^2 L}{\partial x_n^2}
\end{array}\right)$$

where note $\frac{\partial^2 L}{\partial x_i \partial \lambda_j} = -\frac{\partial g_j}{\partial x_i}$ and $\frac{\partial^2 L}{\partial x_i \partial \mu_p} = -\frac{\partial h_p}{\partial x_i}$ for $j = 1, \dots, e$ and $p = 1, .., m$

Conditions for a local constrained max

Let $f(\mathbf{x}), g_1, ..., g_k, h_1, ..., h_m$ be C^2 functions on $\mathbb{R}^n$. Suppose we want to maximize $f(\mathbf{x})$ on the constraint set $C_{g,h} \equiv \{\mathbf{x} = (x_1, ..., x_n) \colon g_1(\mathbf{x}) \leq b_1, ..., g_k(\mathbf{x}) \leq b_k, h_1(\mathbf{x}) = c_1, ..., h_m(\mathbf{x}) = c_m\}$

(1) Form the Lagrangian, then find and solve the FOCs for $(\mathbf{x}^*, \lambda^*, \mu^*)$

(2) Suppose $e \leq k$ inequality constraints are binding at $\mathbf{x}^*$, form the Hessian of L wrt $\mathbf{x}$ at $(\mathbf{x}^*, \lambda^*, \mu^*)$ as on the previous page

- $\rightarrow$ If the bordered Hessian, H, is negative definite on the linear constraint set, then $\mathbf{x}^*$ is a strict local constrained max of f on $C_{g,h}$
- $\rightarrow$ If the last $n - (e + m)$ LPMs of H alternate in sign, with the sign of $|H|$ the same as the sign of $(-1)^n$, then H is negative definite

As we are constrained by the remaining time for Math Camp, we'll maximize the benefit to you by not dwelling on negative semidefiniteness, indefiniteness, or the conditions for a local constrained min. You are left to consider or research them on your own, if you wish

Lab Exercise 1

Exercise: For each of the following functions defined on $\mathbb{R}^2$, find the critical points and classify them as local max, local min, saddle points, or "can't tell"

(1) $x^4 + x^2 - 6xy + 3y^2$

(2) $x^2 - 6xy + 2y^2 + 10x + 2y - 5$

(3) $xy^2 + x^3y - xy$

(4) $3x^4 + 3x^2y - y^3$

Lab Exercise 2

Exercise: Let $Q(K, L) = AK^{\alpha}L^{\beta}$ be a production function faced by a firm, where $A > 0$, $\alpha > 0$, and $\beta > 0$ are constants, and $K \geq 0$ and $L \geq 0$ are real variables chosen by the firm. Let $p = 1$ be the market price the firm receives for selling one unit of output. Let $r > 0$ be the market rental rate (price) of K (capital) and $w > 0$ be the market wage (price) of L (labour) faced by the firm.

(1) What is the profit function of this firm, $\pi(K, L)$?

(2) What are the necessary FOCs for some point (K^*, L^*) in the interior of $\mathbb{R}^2_+$ to be a max of π?

(3) What is Hessian matrix of π?

(4) What are the sufficient SOCs for some interior point (K^*, L^*) to be a max of π?

(5) What are the sufficient SOCs for some interior point (K^*, L^*) to be a unique max of π?

(6) What do these conditions require of the parameters α and β for some interior point (K^*, L^*) to be a unique max of π? What "returns to scale" are exhibited by a production function that satisfies these conditions?

Lab Exercise 3

Suppose we want to maximize a representative consumer's utility, $U(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$, subject to $p_1 x_1 + p_2 x_2 \leq M$, where p_1 and p_2 are the prices of x_1 and x_2, respectively, and M is the money the consumer can spend

(1) Set up the Lagrangian

(2) What are the FOCs. How many are there?

(3) Solve for the critical point(s) of U

(4) Form the bordered Hessian

(5) Check whether the SOCs hold for a critical point to be a constrained max

References

This drew notes from:

- Mathematics for Economists (Simon and Blume, 2015)