Part II, Day 1: Constrained Optimization

Jacob Bradt[1]
jbradt@g.harvard.edu
Harvard Kennedy School

Harvard Public Policy, Health Policy, and Business Administration PhD Math Camp
17 August 2022

[1] These notes are a slight adaptation of previous versions by Robert French, Rebecca Sachs, Todd Gerarden, Wonbin Kang, Sam Richardson, and Jonathan Borck. I am deeply indebted to previous Math Camp TF's for their work on these notes.

Outline

Outline

Introduction

- Fifth-year Public Policy PhD student; undergrad in environmental science/economics at Harvard
- Studying environmental, climate, and energy economics; research includes:
  - Methods used to value non-market environmental amenities
  - Distributional impacts of flooding and flood adaptation investments
  - Clean energy technology policy
- Fields: Environment/Energy, Industrial Organization, Public Finance
- Committee: Joe Aldy, Myrto Kalouptsidi, Ariel Pakes

Roadmap for Part II

- Day 1: Constrained Optimization
  - Substitution/Lagrangian methods, convex/non-convex sets, concavity/quasicocavity

- Day 2:
  1. Comparative Statics
     - Total differentiation/derivatives, implicit functions, comparative statics
  2. Linear Algebra
     - Matrics, matrix operations, matrix properties, matrix inverses, systems of equations

- Day 3: Probability and Statistics
  - Random variables, probability distributions, sampling theory

Goals for Part II

- This is entirely for *your* benefit; you will *not* be graded!
- Provide a refresher of basic concepts, tools, and exercises that will come up often during G1 coursework in microeconomics and econometrics
- Get you back in the practice of solving problems by hand
  - Much of your first (and possibly second) year will be spent on problem sets
- Give you a chance to get to know your peers/classmates
  - Much of your first (and possibly second) year will be spent on problem sets

General structure of Part II

- For a given topic, we will use the following general formula:
  - Brief lecturing on a given topic
  - Interactive example problems, both altogether and in groups
  - If time allows, additional small group problem solving
- Will plan to take hour-long lunch breaks around noon, frequent breaks
- Will likely wrap up each day between 3:00 and 4:00pm
  - I will always stick around during breaks/at the end of the day for questions
- Materials:
  - Notes $\approx$ slides
  - Let $A = \{\text{Example problems in notes/slides}\}$ and $B = \{\text{Problems in exercises handouts}\}$. Then $A \subset B$
  - Will share slides + possible exercise solutions at the end of each day

Final notes before we begin Part II

- This is entirely for *your* benefit, so ask questions throughout!
- I intend to target the median background/exposure, BUT this is *not* intended to be an exhaustive course in all you need to know for your G1 year
  - Our math camp is intended to serve as a refresher
  - I will provide references to texts that may be useful if you feel you need additional resources on a given topic
  - Importantly, your G1 coursework will build up from first principles
- I am not a math teacher
  - I am just someone who has taken G1 coursework in microeconomics (Econ 2120a/b) and econometrics (Econ 2120/2140)
  - There will be things that I can't immediately explain (well)
  - I promise to follow up with an answer if I cannot immediately provide one
- Any questions?

Outline

Some core concepts

- **Unconstrained Optimization**: Find the optimal level of one or more "choice variables" that will either maximize or minimize the "objective function."

- **Constrained Optimization**: Same thing, but we have one or more "constraints" that impose limits on what value(s) our "choice variable(s)" can take.

Constrained optimization examples

- Maximizing a utility function subject to a budget constraint (one equality constraint):

$$\max_{x,y} U(x,y)$$

$$\text{s.t. } p_x x + p_y y = w.$$

  The constraint is a 1-1 mapping between x and y, for a given p_x, p_y and w. It thus reduces the dimensionality of the problem.

- Profit maximizing for a competitive firm (multiple inequality constraints):

$$\max_{x,y} \Pi(x,y)$$

$$\text{s.t. } \Pi(x,y) \geq 0,\ x \geq 0,\ y \geq 0$$

Notation and general approach

- **General notation**:

$$\max_{x,y} f(x,y)$$

$$\text{s.t. } g(x,y) = c$$

→ Sometimes we write this as an implicit equation.

- **General approach**:
  - We work with differentiable functions and use techniques of calculus to solve for optima
  - We will emphasize the use of first order conditions to identify interior optima for optimization problems with equality constraints
  - If time permits, we will discuss second order conditions, boundary solutions, and inequality constraints at a high level, but these will be given an in depth treatment during the semester

Solving constrained optimization problems: The substitution method

The general approach here is to convert a constrained optimization problem into an unconstrained optimization problem:

1. Use the constraint to solve for one variable in terms of the other(s).
2. Substitute the expression from Step 1 into the objective function.
3. Solve this new unconstrained optimization problem as before:
   - Take the first-order condition(s) to find the potential maxima or minima;
   - Check the second-order condition(s) to verify what each candidate solution is; and
   - Take the $\arg\max(\min)$ of the unconstrained function as the $\arg\max(\min)$ of the constrained function.

The substitution method: Example #1

$$\max_{x,y} U(x,y) = 2x + 5\ln y$$

$$\text{s.t. } 6x + 3y = 51$$

We can simplify the constraint to $2x + y = 17$, implying that $2x = 17 - y$.
This gives us $U(x,y) = 17 - y + 5\ln(y)$. We now maximize the one-variable problem, to give:

$$\frac{\partial U}{\partial y} = -1 + \frac{5}{y} = 0 \implies y = 5$$

$$\frac{\partial^2 U}{\partial y^2} = -\frac{5}{y^2} < 0, \forall y$$

$$\implies x = 6$$

The substitution method: Example #2

$$\min_{a,b} C(a,b) = (3a-7)^2 + 4b$$

$$\text{s.t.} \; -24a - 8b = -42$$

We can simplify the constraint to $4b = 21 - 12a$.
This gives us $C = (3a-7)^2 + 21 - 12a$. We now minimize the one-variable problem, to give:

$$\frac{\partial C}{\partial a} = 6(3a-7) - 12 = 0 \implies a = 3$$

$$\frac{\partial^2 C}{\partial a^2} > 0, \forall a$$

$$\implies b = \frac{21-36}{4}$$

Solving constrained optimization problems: The Lagrangian method

Consider the following setup: We have an objective function $z = f(x, y)$ subject to the constraint $g(x, y) = c$, where $c \in \mathbb{R}$ is a constant. To maximize $f()$ subject to the constraint, we have the following steps:

1. Introduce the Lagrange multiplier, λ, and rewrite the constraint with everything on one side of the equation: $g(x, y) - c = 0$.
2. Create the Lagrangian function, a modified version of the objective function

$$\mathcal{L} = f(x, y) - \lambda[g(x, y) - c]$$

3. Solve this unconstrained optimization problem as usual, treating the Lagrange multiplier, λ, as an additional variable.
4. Check your solution from Step 3 to determine if it's a maximum or minimum.

The Lagrangian method: Example #1

Example:

$$\max_{x,y} U(x,y) = 2x + 5\ln y$$

$$\text{s.t. } 6x + 3y = 51$$

Write Lagrangian:

$$\mathcal{L} = 2x + 5\ln y - \lambda(6x + 3y - 51)$$

First order conditions:

$$\frac{\partial \mathcal{L}}{\partial x} = 2 - 6\lambda = 0$$
$$\frac{\partial \mathcal{L}}{\partial y} = \frac{5}{y} - 3\lambda = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = 6x + 3y - 51 = 0$$

Solving the above system gives $y = 5$ and $x = 6$. We further have that $\lambda = \frac{1}{3}$.

The Lagrangian method: Example #2

Example:

$$\max_{x,y} f(x,y) = xy$$
$$\text{s.t. } x + 4y = 16$$

Write Lagrangian:

$$\mathcal{L} = xy - \lambda(x + 4y - 16)$$

First order conditions:

$$\frac{\partial \mathcal{L}}{\partial x} = y - \lambda = 0$$
$$\frac{\partial \mathcal{L}}{\partial y} = x - 4\lambda = 0$$
$$\frac{\partial \mathcal{L}}{\partial \lambda} = x + 4y - 16 = 0$$

Solving the above system gives $y = 2$ and $x = 8$. We further have that $\lambda = 2$.

Interpreting the Lagrange multiplier

- The Lagrange multiplier is often called the "shadow value" or "shadow price" of the constraint
- It expresses how much the objective function changes if we "relax" the constraint a little bit. Or, a measure of the sensitivity of the optimal value of the objective function to changes in the constraint.
- Intuition: penalty on the objective function

Economic interpretation

- *Utility maximization*: The Lagrange multiplier (*when on a budget constraint*) is interpreted as the shadow price of wealth or the marginal utility of wealth
  - → The change in utility that would result from an infinitesimal increase in wealth
- *Profit maximization*: The Lagrange multiplier (*when on the cost function*) for a particular input is interpreted as the shadow price of that input
  - → The change in profits that would result from an infinitesimal increase in use of that good

Deriving an interpretation of λ

Consider the optimization problem,

$$\max_{x,y} f(x,y)$$

$$\text{s.t. } h(x,y) = a$$

Show that,

$$\lambda^*(a) = \frac{d}{da} f(x^*(a), y^*(a)),$$

where x^* and y^* denote the values of x and y that maximize the objective function subject to the constraint.[2]

[2] Hint: it may be useful to define the function $F(a) = h(x^\star(a), y^\star(a)) - a$, and note that $\frac{\partial F}{\partial a}$ will always equal zero.

Deriving an interpretation of λ

Start with the R.H.S. and use the chain rule.

$$
\begin{aligned}
\frac{d}{da} f(x^*(a), y^*(a)) &= \frac{\partial f}{\partial x^*}\frac{\partial x^*}{\partial a} + \frac{\partial f}{\partial y^*}\frac{\partial y^*}{\partial a} \\
&= \lambda \frac{\partial h}{\partial x^*}\frac{\partial x^*}{\partial a} + \lambda \frac{\partial h}{\partial y^*}\frac{\partial y^*}{\partial a} \\
&= \lambda \left(\frac{\partial h}{\partial x^*}\frac{\partial x^*}{\partial a} + \frac{\partial h}{\partial y^*}\frac{\partial y^*}{\partial a} \right) \\
&= \lambda \left(\frac{\partial h(x^*(a), y^*(a))}{\partial a} \right) \\
&= \lambda
\end{aligned}
$$

where the last equality follows because $h(x^*(a), y^*(a)) = a \Longrightarrow \frac{\partial h(x^*(a), y^*(a))}{\partial a} = 1$.

Inequality constraints

- Many problems in economics have either (1) non-binding constraints, or (2) the possibility of corner solutions

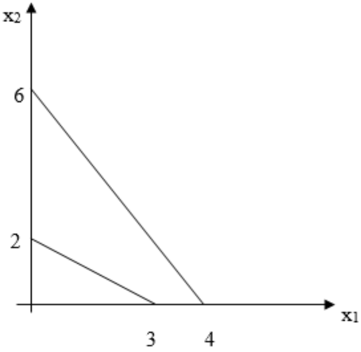
- Examples of non-binding constraints include:

1. Non-negativity constraints on production inputs, e.g.:

$$\min_{x_1, x_2} C = (x_1 - 4)^2 + (x_2 - 4)^2$$

s.t.

$$\begin{aligned} 2x_1 + 3x_2 &\geq 6 \\ -3x_1 - 2x_2 &\geq -12 \\ x_1, x_2 &\geq 0 \end{aligned}$$



2. Non-negativity constraints on firm profits.

Inequality constraints

- Many problems in economics have either (1) non-binding constraints, or (2) the possibility of corner solutions

- Examples of corner solutions include:
  - Consumer demand for a subset of available goods;
  - Production using a subset of available goods.

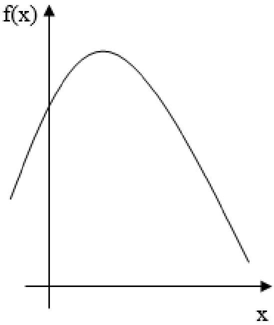


Figure 1.

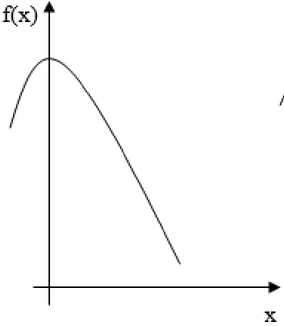


Figure 2.

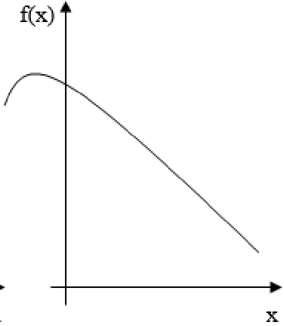


Figure 3.

Solving constrained optimization problems with inequality constraints

- In a lot of cases, the best way to deal with non-binding constraints is to solve the problem ignoring the constraints and ex-post check that the solutions satisfy these constraints
- If the solutions do not satisfy these constraints, we can use Kuhn-Tucker conditions
  - $\rightarrow$ Not covering today, but will introduce in microeconomics sequence and include a supplemental section in the Part II notes covering Kuhn-Tucker conditions with a worked example

Sets

- A **set** is a collection of objects (often called elements). These objects may indeed be numbers.
- *Examples*:
  - In one-dimensional Euclidean space, a line segment or series of line segments:

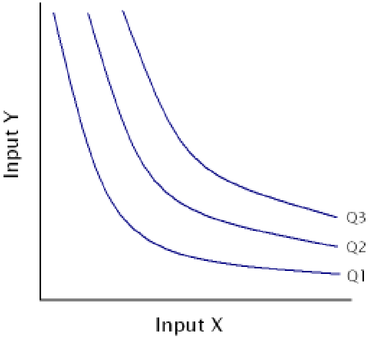
    $$(0,1); \{0,1\}; \{(0,1), 1, [1,3)\}$$

    Note here how intervals are either defined as a set or an element in a set.
  - This notion of sets applies to higher dimensional Euclidean space. Moreover, sets do not have to contain elements in Euclidean space; here is a set with three objects, for example:

    $$\{\text{red, white, blue}\}$$

Application of sets: Level sets

- One (of many) ways in which sets and set notation often comes up is in the context of **level sets**
  - Used to study two fundamental functions of microeconomics: production and utility functions
- Level sets provide an intuitive way of understanding a function that maps from $\mathbb{R}^n \to \mathbb{R}^1$
  - Describe all combinations of n inputs that produce a given function value
- *Example:* simple Cobb-Douglas production function $Q = f(x, y) = x \cdot y$ where x and y measure amounts of two inputs and Q is output



Convex sets

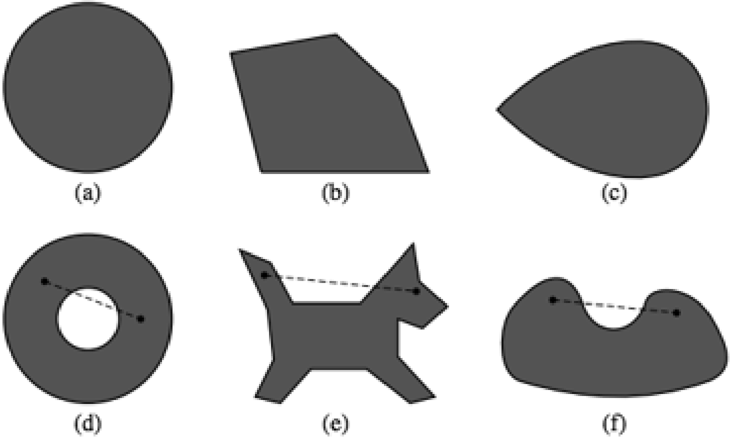
- A **convex set** in Euclidean space is a set $\in \mathbb{R}^n$ where the line segment joining any two points in the set is contained entirely within the set.

- Algebraically, a set, call it C, is convex if and only if $\forall t \in [0, 1]$, and $\forall x, y \in C$, we have that
$$tx + (1 - t)y \in C$$

- If a set C does not satisfy the above condition, we call it a **non-convex set**

Convex sets

Examples:



Concave and convex functions

- **Concave function:** A real-valued function f defined on a convex subset $U \subset \mathbb{R}^n$ is concave if $\forall \mathbf{x}, \mathbf{y} \in U$ and $t \in [0, 1]$

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \geq tf(\mathbf{x}) + (1-t)f(\mathbf{y})$$

- **Convex function:** A real-valued function g defined on a convex subset $U \subset \mathbb{R}^n$ is convex if $\forall \mathbf{x}, \mathbf{y} \in U$ and $t \in [0, 1]$

$$g(t\mathbf{x} + (1-t)\mathbf{y}) \leq tg(\mathbf{x}) + (1-t)g(\mathbf{y})$$

Concave and convex functions

- Standard approach you've likely seen: can tell whether or not a function on $\mathbb{R}^n$ is concave by looking at its graph in $\mathbb{R}^{n+1}$

- This is challenging in higher dimensions! Thus, we have the calculus criteria, which you're likely familiar with in the single dimension case:
  - Let f be a C^1 function on a convex subset U of $\mathbb{R}^n$. Then f is concave on u iff forall $\mathbf{x}, \mathbf{y} \in U : f(\mathbf{y}) - f(\mathbf{x}) \leq Df(\mathbf{x})(\mathbf{y} - \mathbf{x})$

- Given their appeal for higher order problems (easy to graph for $f : \mathbb{R}^2 \rightarrow \mathbb{R}$) and their nice economic intuition, we will often work with level sets $\Longrightarrow$ develop a definition/intuition for concavity/convexity by looking at a function's level set

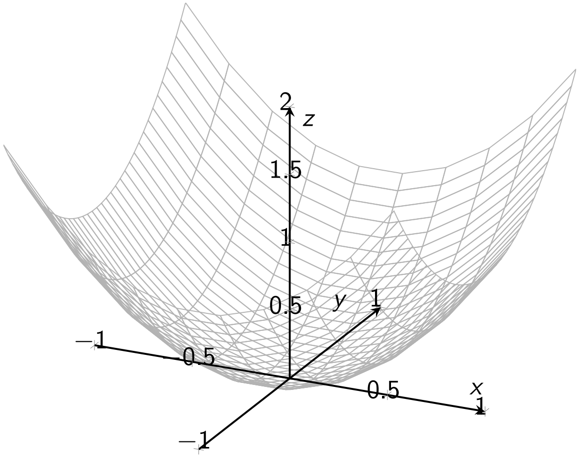
- Let f be a function defined on a convex set $U \subset \mathbb{R}^n$. If f is concave then for every $\mathbf{x_0} \in U$, the set

$$C^+_{\mathbf{x_0}} \equiv \{\mathbf{x} \in U : f(\mathbf{x}) \geq f(\mathbf{x_0})\}$$

  is a convex set

Concave and convex functions

Example: is the function $z = f(x, y) = x^2 + y^2$ concave or convex?



What if we look at the level set with $z = 4$?

Properties of concave and convex functions

Why do we care about concavity/convexity?

- Let f be a concave (convex) function defined on $U \subset \mathbb{R}^n$. If $\mathbf{x}^*$ is a critical point of f then $\mathbf{x}^* \in U$ is a global maximizer (minimizer) of f on U.

- Let $f_1, \ldots, f_k$ be concave (convex) functions each defined on the same subset $U \subset \mathbb{R}^n$ and let $a_1, \ldots, a_k > 0$. Then $a_1 f_1 + \ldots + a_k f_k$ is a concave (convex) function on U.

Concave functions in economics

- Expenditure and cost functions are concave
  - Expenditure function:

$$e(p, u) = \min\{p_1 x_1 + ... + p_n x_n : u(x) \geq u\}$$

  - Cost function:

$$c(w, y) = \min\{w_1 x_1 + ... + w_n x_n : g(x) = y\}$$

- Properties of concave (convex) functions are very useful; however, concave functions have a clear downside in economic analysis: concavity is a **cardinal** property
  - → It depends on the numbers which the function assigns to the level sets, not just on the shape of the level sets
  - → In other words, a monotonic transformation of a concave function need not be concave

Monotonic transformations

- We typically apply monotonic transformations to convert difficult-to-analyze functions into easy-to-analyze functions with exactly the same optima
  $\rightarrow$ Also related to ordinal vs. cardinal distinction

- A **positive (negative) monotonic function** is a function that increases (decreases) throughout its domain
  - A positive (negative) monotonic function can be either strictly increasing (decreasing) or non-decreasing (non-increasing)
  - Algebraically, a non-decreasing monotonic function has the property that $\forall x, y$ such that $x \leq y \Rightarrow f(x) \leq f(y)$ (reverse for non-increasing monotonic function)
  - Replacing these inequalities with strict inequalities yields the definition of a strictly increasing monotonic function

- A **monotonic transformation** is achieved by plugging the function you want to analyze into any monotonic function of your choice

- Key result: Any monotonic transformation of a function has the same optima as the original function!

Monotonic transformations: Economic application

Example: The Cobb-Douglas utility function has the form

$$u(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}.$$

Check that by taking the natural log of the function (i.e. take $\ln\left(u(x_1, x_2)\right)$), the optima of the two functions are the same under the constraint $x_1 + x_2 \leq 100$.

Lagrangian with the monotonic transformation:

$$\mathcal{L} = \alpha \ln(x_1) + (1-\alpha)\ln(x_2) - \lambda(x_1 + x_2 - 100)$$

First order conditions:

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\alpha}{x_1} - \lambda = 0 \qquad \frac{\partial \mathcal{L}}{\partial y} = \frac{1-\alpha}{x_2} - \lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -(x_1 + x_2 - 100) = 0$$

Solving gives $x_1 = 100\alpha$, $x_2 = 100(1-\alpha)$ and $\lambda = \frac{1}{100}$

Monotonic transformations: Economic application

Example: The Cobb-Douglas utility function has the form

$$u(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}.$$

Check that by taking the natural log of the function (i.e. take $\ln\left(u(x_1, x_2)\right)$), the optima of the two functions are the same under the constraint $x_1 + x_2 \leq 100$.

Lagrangian without the monotonic transformation:

$$\mathcal{L} = x_1^{\alpha} x_2^{1-\alpha} - \lambda(x_1 + x_2 - 100)$$

First order conditions:

$$\frac{\partial \mathcal{L}}{\partial x} = \alpha x_1^{\alpha-1} x_2^{1-\alpha} - \lambda = 0 \qquad \frac{\partial \mathcal{L}}{\partial y} = -(\alpha - 1) x_1^{\alpha} x_2^{-\alpha} - \lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -(x_1 + x_2 - 100) = 0 \qquad x_1 + x_2 \leq 100$$

Messy!

Monotonic transformations: Econometric application

Example: The likelihood function takes the form,

$$\mathcal{L}(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta),$$

and the loglikelihood takes the form

$$\ell(\theta|x) = \log \prod_{i=1}^{n} f(x_i|\theta) = \sum_{i=1}^{n} \log f(x_i|\theta).$$

Quite useful: is useful multiplying small values makes very small values, and differentiation is easier when the function is additive

Monotonic transformations: Another example!

Test yourself again:

$$\max_{x_1, x_2} f(x) = e^{\sqrt{x_1 x_2}} \text{ s.t. } x_1 + 4x_2 = 16.$$

Lagrangian without monotonic transformation:

$$\mathcal{L} = e^{\sqrt{x_1 x_2}} - \lambda(x_1 + 4x_2 - 16)$$

versus Lagrangian with monotonic transformation $\ln(f(x))$:

$$\mathcal{L} = \sqrt{x_1 x_2} - \lambda(x_1 + 4x_2 - 16)$$

One is *way* easier than the other!

Cardinal vs. ordinal

- A characteristic of functions is called **ordinal** if every monotonic transformation of a function with this characteristic still has this characteristic
- **Cardinal** properties are not preserved by monotonic transformations
- Importantly, utility is an ordinal concept
  - For example, let $u(x, y) \in \mathbb{R}^2_+$ be a utility function and let $v(x, y) = u(x, y) + 1$ be another utility function $\Rightarrow$ same set of indifference curves $\Rightarrow$ same preferences
  - Concavity/convexity desirable properties not applicable when working with utility functions because they are cardinal

Where are we going?

- Concave functions have one fundamental ordinal property: their level sets bound convex sets from below

- It turns out that this property is quite useful and we define a class of functions which have this desired ordinal property of concave functions: quasiconcave (quasiconvex)

- Final definition before introducing quasiconcavity/quasiconvexity: let f be a function defined on the subset $S \subset \mathbb{R}^n$
  - **Upper level set:** for any $a \in \mathbb{R}$

$$P_a^+ \equiv \{x \in S : f(x) \geq a\}$$

  - **Lower level set:** for any $a \in \mathbb{R}$

$$P_a^- \equiv \{x \in S : f(x) \leq a\}$$

Quasiconcavity

- The function f of many variables defined on a convex set S is **quasiconcave** if every upper level set of f is convex
  - That is, $P_a^+ = \{x \in S : f(x) \geq a\}$ is convex for every value of a

- We also have an equivalent algebraic representation: A function f is **quasiconcave** if and only if, for every pair of distinct points u and v in the domain of f, and for $\theta \in (0, 1)$,

$$f(v) \geq f(u) \implies f(\theta u + (1 - \theta)v) \geq f(u).$$

- If the second inequality is strict, then f is **strictly quasiconcave**.

Quasiconvexity

- The function f of many variables defined on a convex set S is **quasiconvex** if every lower level set of f is convex
  - That is, $P_a^- = \{x \in S : f(x) \leq a\}$ is convex for every value of a

Checking quasiconvexity and quasiconcavity

- To see whether a function is quasiconcave or quasiconvex one can examine the level sets of the function directly
- Alternatively if the function is differentiable (twice differentiable in one case), two helpful propositions can determine quasi-concavity and quasiconvexity.[3]. We present them in the supplementary section on optimization, since they do not provide much intuition and require material from later chapters of these notes.

[3]For more information, see Osborne, chapter 3.4 and the cites therein.

Useful properties of quasiconcavity/quasiconvexity

Unlike concavity and convexity, quasi-concavity and quasi-convexity retain their properties of quasiconcavity/quasiconvexity when they are monotonically transformed, which is a useful property for certain objective functions (utility functions) to have

Properties:

- Every concave (convex) function is quasiconcave (quasiconvex)
  - The converse is not necessarily true
- If $f(x)$ is quasiconcave, then $-f(x)$ is quasiconvex
- Any monotonic transformation of a quasiconcave (quasiconvex) function is also quasiconcave (quasiconvex).

Key result

- Knowing whether a function is *strictly* quasiconcave or *strictly* quasiconvex implies that any local optima are also global optima
  - There is thus no need to check second-order conditions if $f(x)$ is strictly quasiconcave or strictly quasiconvex, for finding the FOC of a strictly quasiconcave (strictly quasiconvex) function finds a global maximum (minimum)

- Strictly quasiconcave $\implies$ local optimum $=$ global maximum

- Strictly quasiconvex $\implies$ local optimum $=$ global minimum

Quasiconcavity/quasiconvexity example

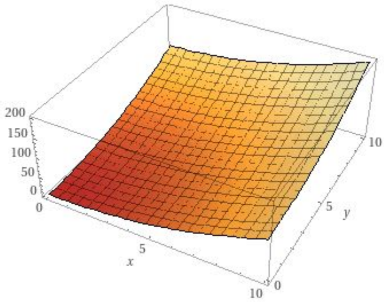
Example: Find the extremum of

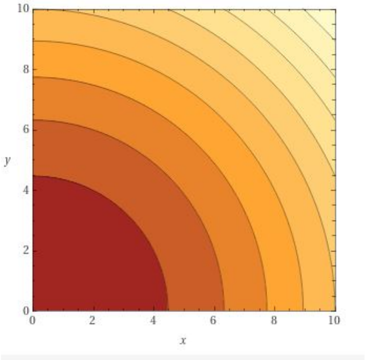
$$f(x_1, x_2) = x_1^2 + x_2^2$$

subject to,

$$x_1 + 4x_2 = 2.$$

$$x_1, x_2 \geq 0$$





Quasiconcavity/quasiconvexity example

Example: Find the extremum of

$$f(x_1, x_2) = x_1^2 + x_2^2$$

subject to,

$$x_1 + 4x_2 = 2.$$

$$x_1, x_2 \geq 0$$

Level sets are strictly quasiconvex $\implies$ optimum is a global minimum. The Lagrangian is

$$\mathcal{L} = x_1^2 + x_2^2 - \lambda(x_1 + 4x_2 - 2)$$

The FOCs are

$$\frac{\partial \mathcal{L}}{\partial x_1} = 2x_1 - \lambda = 0 \qquad \frac{\partial \mathcal{L}}{\partial x_2} = 2x_2 - 4\lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = x_1 + 4x_2 - 2 = 0$$

The first two equations give us $x_1 = \frac{x_2}{4}$, and substituting this into the third equation we get $x_2 = \frac{8}{17}$ and $x_1 = \frac{8}{4 \cdot 17}$, with $\lambda = \frac{4}{4 \cdot 17}$.

Constrained optimization with multiple constraints

This is identical to the case with a single constraint, aside from adding an additional Lagrange multiplier for each constraint. Note though, you can still substitute in constraints where possible.

Constrained optimization with multiple constraints

Consider an objective function $z = f(x, y)$ subject to two constraints, $g(x, y) = c$ and $h(x, y) = d$:

1. Introduce two Lagrange multipliers, λ_1 and λ_2, one for each constraint;
2. Rewrite each constraint with everything on one side of the equation:

$$g(x, y) - c = 0 \text{ and } h(x, y) - d = 0$$

3. Create the Lagrangian function, a modified version of the objective function:

$$\mathcal{L} = f(x, y) - \lambda_1 [g(x, y) - c] - \lambda_2 [h(x, y) - d]$$

4. Solve this unconstrained optimization problem as usual, treating the Lagrange multipliers, λ_1 and λ_2, as additional variables.
5. Check your solution from Step 4 to determine if it's a maximum or minimum. You can use the bordered Hessian approach outlined in Simon and Blume, chapter 19.

Constrained optimization with multiple constraints: Example

Example:

Find the extremum of $z = x^2 + 2xy + yw^2$ subject to

$$2x + y + w^2 = 24$$

and

$$x + w = 8.$$

The Langrangian is

$$\mathcal{L} = x^2 + 2xy + yw^2 - \lambda(2x + y + w^2 - 24) - \mu(x + w - 8)$$

The FOCs are:

$$\frac{\partial \mathcal{L}}{\partial x} = 2x + 2y - 2\lambda - \mu = 0 \qquad \frac{\partial \mathcal{L}}{\partial y} = 2x + w^2 - \lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial w} = 2yw - 2\lambda w - \mu = 0 \qquad \frac{\partial \mathcal{L}}{\partial \lambda} = 2x + y + w^2 - 24 = 0$$

$$\frac{\partial \mathcal{L}}{\partial \mu} = x + w - 8 = 0$$

Key takeaways

- Know solution methods (Lagrangian) for constrained optimization problem with single/multiple equality/inequality constraints
- Be familiar with sets/set notation, level sets, definition of convex/non-convex sets
- Be familiar with quasiconcavity/quasiconvexity *and the implications for constrained optimization*

Additional resources

- Chiang and Wainwright, chapter 12
- Simon and Blume, chapters 18-19, 21
- Martin Osborne's economic math *website*.