Math Review
Summer 2016

Topic 8

8. Optimization

8.1 Unconstrained maximization

We start this lesson by considering the simplest of optimization problems, those without conditions, or what we refer to as **unconstrained optimization problems**. Unconstrained optimization is equivalent to finding the max/min of a function. In one-variable calculus we were able to find these points by finding the critical points of a function and looking at the sign of the second derivative at that point. In multivariable calculus the results are much the same. To identify maxima and minima we again need to be able to find the critical points of f:The point $\hat{x} \in \mathbb{R}^n$ is a critical point of $f: \mathbb{R}^n \to \mathbb{R}$ if $Df(\hat{x}) = 0$.

Recall 2 things from our previous classes:

(1) The local maximum, x^*, can be either a boundary max or an interior max. If x^* is an interior max, then x^* is a critical point.
(2) If the function f has a critical point in $C \subset \mathbb{R}^n$ then the Hessian matrix can be used to identify whether the critical point is a maximum or minimum.

Let's remember those cases:

$Df(x^*)$	$D^2f(x^*)$	Max/Min
$= 0$	Negative semidefinite	?
$= 0$	Positive semidefinite	?
$= 0$	Neither	?

- The $n \times n$ matrix $\boldsymbol{A}$ is positive definite if and only if its n principal minors are all greater than 0: $detA_1 > 0, detA_2 > 0, \ldots, detA_n > 0$.
- The $n \times n$ matrix $\boldsymbol{A}$ is negative definite if and only if its n principal minors alternate in sign with the odd order ones being negative and the even order ones being positive: $det\ A_1 < 0, det\ A_2 > 0, detA_3 < 0, detA_4 > 0, \ldots$.

- The $n \times n$ matrix $\boldsymbol{A}$ is positive semidefinite if and only if its principal minors are all greater than or equal to 0: $det\ A_1\ \geq\ 0\ and\ det\ A_2\ \geq\ 0, \ldots, A_n\ \geq\ 0.$
- The $n \times n$ matrix $\boldsymbol{A}$ is negative semidefinite if and only if its n principal minors alternate in sign with the odd order ones being less than or equal to zero and the even order ones being greater than or equal to 0: $det\ A_1\ \leq\ 0\ and\ det\ A_2\ \geq\ 0, \ldots.$

Example. Consider this example.

$$f(x,y) = x^4 + x^2 - 6xy + 3y^2$$

Let's find the critical points and classify them as local max, local min or saddle points.

$\mathcal{Q}$: Solve the following problem: A monopolist producing a single output has two types of customers. If it produces q_1 units for type 1, then these customers are willing to pay a price of $50 - 5q_1$ per unit. If it produces q_2 units for type 2, then these customers are willing to pay a price of $100 - 10q_2$ per unit. The monopolist's total cost of manufacturing q units of output is $90 + 20q$.

Set up this monopolist profit function given by $\pi(q_1, q_2)$. In order to maximize profits, how much should the monopolist produce for each market? Can you ensure that this production plan yields a maximum?

You should be able to read a statement as above and frame that in a maximization/minimization problem. Try your best and we'll help each other.

$\mathcal{A}$:

8.2 Kuhn Tucker and constrained maximization

8.2.1 General forms of optimization

While you will work through a few unconstrained max/min problems, most economic decisions are the result of an optimization problem subject to one or a series of constraints:

- Consumers make decisions on what to buy constrained by the fact that their choice must be affordable.
- Firms make production decisions to maximize their profits subject to the constraint that they have limited production capacity.
- Households make decisions on how much to work/play with the constraint that there are only so many hours in the day.
- Firms minimize costs subject to the constraint that they have orders to fulfill.

All of these problem fall under the category of constrained optimization and luckily for us, there is a uniform process that we can use to solve these problems. {BUS}

Constrained optimization problems are of the general form:

$$\max_{\{x_1,x_2,\ldots,x_n\}} f(x_1,x_2,\ldots,x_n)$$

subject to:

$$g_1(x_1,x_2,\ldots,x_n) \leq b_1$$
$$\vdots$$
$$g_k(x_1,x_2,\ldots,x_n) \leq b_k$$

$$h_1(x_1,x_2,\ldots,x_n) = c_1$$
$$\vdots$$
$$h_m(x_1,x_2,\ldots,x_n) = c_m$$

We call $f(x_1,x_2,\ldots,x_n)$ the objective function and $g_1,g_2,\ldots,g_k$ and $h_1,h_2,\ldots h_m$ are the constraint functions. The constraints $g_i(x_1,x_2,\ldots,x_n) \leq b_i$ for $i = 1,2,\ldots,k$ are the inequality constraints and the $h_j(x_1,x_2,\ldots,x_n) = c_j$ for $j = 1,2,\ldots,m$ are the equality constraints.

A standard consumer problem may take the following form. A consumer with budget I wishes to purchase the goods $x_1, x_2, \ldots x_n$ with prices $p_1, p_2, \ldots p_n$ so as to maximize utility, $u(x_1, x_2, \ldots x_n)$. The consumer's problem is then:

The inequality constraints $x_i \geq 0$ are usually referred to as non-negativity constraints as they restrict the consumer from purchasing a negative amount of some good i.

8.2.2. The Lagrangian/Langragean (*toMAYto-toMAHto*)

Let's start simple and build up to more involved optimization problem.

You are given this maximization problem framed as:

$$\max_{\{x_1,x_2\}} f(x_1, x_2)$$
$$\text{subject to:}$$
$$h(x_1, x_2) = c$$

There are two ways of solving the above maximization problem with given constraints. The first via *substitution*, the second via a *Lagrangean*. When dealing with constrained optimization, you often want to form a Lagrangean denoted by $\mathcal{L}$. You may read about the substitution method but most of Micro will really revolve around building the $\mathcal{L}$. At first this method may seem more difficult, but which in fact can be very useful (and sometimes easier). I will go through the formal definition, but the application is even simpler.

Lagrange's Theorem: Let $f(x)$ and $g_j\ (x), j = 1, \ldots, m$ be a continuously differentiable real-valued functions over some domain $D \to \mathbb{R}^n$. Let x^* be an interior point of D and suppose that x^* is an optimum (maximum or minimum) of f subject to the constraints, $g_j\ (x) = 0$. If the gradient vectors $\nabla g_j\ (x^*), j = 1, \ldots, m$, are linearly independient, then there exists m unique numbers $\lambda_j^*, j\ =\ 1, \ldots, m$, such that:

$$\frac{\partial L(x^*, \lambda^*)}{\partial x_i} = \frac{f(x^*)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j(x^*)}{\partial x_i} = 0$$

For $i = 1,2, \dots, n$. $L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x^*)$

There is a lot at work here that we are not touching on. For example, we require that $\frac{\partial g}{\partial x_i} \neq 0, \forall i = 1,2, .., m.$ But in Micro, you will usually form the Lagrangean and not really have to worry too much about this. If these partials go to zero, it will be very obvious. In a nutshell here, you want to form:

$\mathcal{L}(\mathrm{x}_1, \mathrm{x}_2, \lambda) = f(x_1, x_2) + \lambda(c - h(x_1, x_2))$ from our example above.

Or equivalently:

$\mathcal{L}(\mathrm{x}_1, \mathrm{x}_2, \lambda) = f(x_1, x_2) - \lambda(h(x_1, x_2) - c)$

The first order conditions are:

These first order conditions can then be used to solve for x_1, x_2, λ

It may be valuable to note that: $\lambda^* = \frac{\frac{\partial f}{\partial x_1}(x_1^*, x_2^*)}{\frac{\partial h}{\partial x_1}(x_1^*, x_2^*)} = \frac{\frac{\partial f}{\partial x_2}(x_1^*, x_2^*)}{\frac{\partial h}{\partial x_2}(x_1^*, x_2^*)}$. - (L)

Working with the first and second equation often will eliminate the λ^*, which is beauty of this method to problem solve. λ^* can then be solved for by plugging the results back into the last ugly equation (L).

Group example. Let's check how good our skill of applying definitions work: You are given a problem framed as follows:

$$\max_{\{x_1,x_2\}} -ax_1^2 - bx_2^2$$

subject to:

$$x_1 + x_2 = 1$$

Form the Lagrangean, write out the FOCs and solve.

$\mathcal{Q}$: Now try this maximization problem on your own:

$$\max_{x_1,x_2} x_1^2 x_2$$
subject to
$$2x_1^2 + x_2^2 = 3$$

$\mathcal{A}$:

Economic interpretation of the Lagrangian:

In economics, the values of λ^* often have important economic interpretations. Consider a budget constraint given by $p_1x_1 + p_2x_2 = w$, that is price multiplied by commodities add up to your wealth. Now think of a more general constraint of the form $g_j(x_j) = b_j$. If the right hand side b_j of constraint j is increased by Δ, say by one unit, then the optimum objective value increases by approximately $\lambda_i^*\Delta$ (or simply λ_i^* if Δ is one unit) . We won't expand too much on that. Often the λ_i^* simply drops out, but it is valuable to know where it all is coming from. In our consumer problem, with the budget constraint, parameters like wealth are usually fixed, but really, all of our solutions are dependent on that value. In reality, we can denote $x_1^*(w), x_2^*(w), \lambda^*(w)$ as the solutions. As wealth changes, these will change. You will find that Prof. Glewwe calls this the shadow value of wealth in the consumer problem.

Try at home:

Solve for x_1 and x_2

$\min_{x_1,x_2} x_1^2 + x_2^2$, subject to $x_1x_2 = 1$

$\min_{x_1,x_2} x_1x_2$, subject to $x_1^2 + x_2^2 = 1$

$\max_{x_1,x_2} x_1x_2^2$, subject to $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$

How about a general consumer problem? Solve for x, y

$\max_{x,y} kx^\alpha y^\beta$, subject to $p_x x + p_y y = m$

Now a general producer problem, solve for K, L

$\max_{K,L} wL + rK$, subject to $AK^\alpha L^\beta = q$

8.2.2. Kuhn-Tucker and inequality conditions

We can really spend a whole day on Kuhn-Tucker conditions – but we won't! Let's go through the essential material you will need to know in order to be able to keep up with Micro theory. **I do not expect you to understand this fully today, it took me many weeks to get it when I was first exposed**. However, let's try our best to get the most out of this section. Let's try not to get too stuck on details – if not today, you WILL get it over the next few times.

In economics it is much more common to start with inequality constraints of the form $g\ (x, y)\ \leq\ c$, for instance in our consumer problem, we would have, $p_1 x_1 + p_2 x_2 \leq w$. The constraint is said to be **binding** if at the optimum $g\ (x^*, y^*)\ =\ c$, and it is said to be slack otherwise. Luckily for us, with a small tweak, the Lagrangean can still be used with the same FOC's except now we have **three "Kuhn-Tucker" necessary conditions** for each inequality constraint.

$$\frac{\partial \boldsymbol{\mathcal{L}}}{\partial \boldsymbol{\lambda}} = c - g\ (x^*, y^*) \geq 0$$

$$\boldsymbol{\lambda}^* \geq \mathbf{0}$$

$$\frac{\partial \boldsymbol{\mathcal{L}}}{\partial \boldsymbol{\lambda}} \boldsymbol{\lambda}^* = \lambda^*[c - g\ (x^*, y^*)] = 0$$

Proof of Kuhn-Tucker can be found on pp. 959-960 Mas-Colell.

A few things to note:

- The first condition is just a restatement of the constraint.
- The second condition says that λ^* is always non-negative.
- The third condition says that either λ^*or $c - g\ (x^*, y^*)$ must be zero.
- If $\lambda^* = 0$, intuitively, we are not giving any weight to the constraint. The problem can really turn into an unconstrained optimization problem.
- If $\lambda^* > 0$, then the constraint must be binding then the problem turns into the standard Lagrangean considered above.

All in all, the Kuhn-Tucker condition provide a neat mathematical way of turning the problem into either an unconstrained problem or a constrained one. Unfortunately you **typically have to check both cases**, unless you know for sure it's one or the other. Let's work through an example together to get at what the Kuhn-Tucker does.

Example.

$$\max_{x_1,x_2} u(x_1, x_2) = 4x_1 + 3x_2$$

subject to:

$$g(x_1, x_2) = 2x_1 + 3x_2 \leq 10$$

$x_1, x_2 \geq 0$, non-negativity constraints

It would be cruel of me to have you solve a full blown Kuhn-Tucker problem today. I do want you to be able to read them and understand what you are looking at. These are general examples from the Production mini.

This is a profit maximization problem framed as:

$$\max_{q\geq 0, z\geq 0} \pi(p,r) = p.q - r.z$$
$$\text{subject to:}$$
$$D_o(q,z) \leq 1$$

I will go over what the different components mean briefly.

How can we set up the Lagrangean?

Exercise and example. Let's have you solve one of Prof. Terry Hurley famous profit maximization problems. If you can do this – you are at least 20% ready to take a Production final ☺

Consider this profit maximization problem with two outputs and two inputs constrained by an input distance function given by $D_1(q,z)$.

$$\max_{q\geq 0, z\geq 0} \pi(p,r) = p_1q_1 + p_2q_2 - r_1z_1 - r_2z_2$$

subject to:

$$D_I(q,z) = \frac{\sqrt{z_1z_2}}{q_1^2 + q_2^2} \geq 1$$

Form the Lagrangean for this problem:

Write out the first order conditions (with Kuhn-Tucker):

~~Let's not worry about checking for Kuhn-Tucker conditions and find all the solutions in general form, that is, solve for $z_1, z_2, q_1,$and q_2 given they are interior solutions. Before jumping into solving, remember that we want to find our results in the form of: $z(p,r)$ and $q(p,r)$.~~

This took me over 3 pages of algebra to get a presentable answer. Now, **that** would be cruel to have the class go through. Please try at home if desired. For now, let's do a simplier example from the production mini itself which covers all the bases.

Example. Consider this cost minimization problem two inputs constrained by an input distance function given by $D_1(q,z)$.

$$\min_{z\geq 0} C = r_1 z_1 - r_2 z_2$$

subject to:

$$D_I(q,z) = \frac{\sqrt{z_1 z_2}}{q_1^2 + q_2^2} \geq 1$$

Write down the Lagrangian, Kuhn-Tucker conditions and solve for $z_1(r,q)$ and $z_2(r,q)$ assuming interior solutions afterwards. Find the cost function $c(r,q)$ as well.

Morning Exercise.

This is a very simple example to get your brains in the groove this morning.

You are given the utility function below:

$$U = xy$$

Suppose the budget information is as follows: wealth = 100, $p_x = 1$, $p_y = 1$.

1. Write down this person's budget constraint.
2. Set up the utility maximization problem
3. Write down the Lagrangean
4. Solve for x and y assuming interior solutions.

Optimizing with expected values

You will run into some optimization with expected values in consumer theory and game theory. We will do 1 example together and move on. Consider this part of a question:

Suppose there are 2 firms in a market, Firm 1 and Firm 2. We want to solve for the Bayesian Nash equilibrium in a duopoly setting where firm 2 has entered the market and believes that there is a certain probability that firm 1 has high costs and a certain probability that firm 1 has low costs. We want to solve for the expected level of profit for firm 2. The marginal cost for Firm 2 is: 12.

Assume the inverse market demand curve is given by $P(Q) = 24 - Q$, where P is price and $Q = q_1 + q_2$ is market quantity produced.

Write out Firm's 2 problem:

Find the FOC:

Suppose, you further know that Firm 1 has costs of $c_i(q_1) = c_i q_1$, where q_1 is the quantity produced by firm 1, and marginal cost, c_i, equals c_H or c_L. Assume that high marginal cost are $c_H = 12$ and low marginal cost are $c_L = 6$. Firm 2 can assume high costs occur with probability 0.5 and low costs occur with probability 0.5. Solve for q_2.

<u>Optimization with 'odd' forms</u>:

Complements

You are given a utility function:

$$\max_{x,y} u(x,y) = \min\,(2x, y)$$
$$\text{subject to}$$
$$p_1x + p_2y = m$$

We start by noting that this utility function is not differentiable at the kink $2x = y$. The optimal allocation (x, y) must satisfy $2x = y$ since it would be the cheapest way to achieve.

$\mathcal{Q}$: This should be very similar, just to give you a tiny bit of practice. You have a utility function given by: $u(x_1, x_2) = \min\{\alpha x_1, \beta x_2\}$. The price of good 1 is p_1 and price of good 2 is p_2, with wealth, w. Solve for this consumer problem, find x_1 and x_2.

$\mathcal{A}$:

Substitutes

Suppose that Jack's utility is entirely based on number of hours spent camping (c) and skiing (s). You are given his maximization problem denoted as:

$$\max_{c,s} u(c,s) = 3c + 2s$$
$$\text{s.t}$$
$$p_c c + p_s s = m$$

Form the Lagrangian and solve for c and s.

Q: Can you try with the following example?

You are given this maximization problem denoted as:

$$\max_{x_1,x_2} u(x_1,x_2) = 3x_1 + 5x_2$$
$$\text{s.t}$$
$$p_1x_1 + p_2x_2 = m$$

8.3. Envelope Theorem

Sometimes, the objective function f is also a function of some of the exogenous parameters, for example, i.e. $f(x;\ a)$. The envelope theorem is a general principle describing how the value of an optimization problem changes as the parameters of the problem change. In other words, we are looking at how the maximal value of a function depends on some parameters. You will most likely see envelope theorem applied in 'dual' problem. For example, cost minimization is a dual problem to utility maximization for a consumer.

For formality:

Envelope Theorem: Let $f, h_1, h_2 \dots, h_k$ be continuously differentiable functions. Let $x^*(a) = (x_1^*(a), x_2^*(a), \dots, x_3^*(a))$ be the solution to:

$$\max_{\{x_1, x_2, \dots, x_n\}} f(x_1, x_2, \dots, x_n; a)$$

subject to:

$$h_1(x_1, x_2, \dots, x_n; a) = 0$$
$$h_2(x_1, x_2, \dots, x_n; a) = 0$$
$$\vdots$$
$$h_k(x_1, x_2, \dots, x_n; a) = 0$$

For any fixed choice of the exogenous parameter a. Suppose that the Jacobian matrix associated with the equality constraints has maximum rank. If $x^*(a), \lambda_1(a), \dots \lambda_k(a)$ are all continuously differentiable functions of a, then:

$$\frac{\partial f(x^*(a); a)}{\partial a} = \frac{\partial \mathcal{L}}{\partial a}(x^*(a), \lambda(a); a)$$

Think of this 'dual' problem:

$$\min wL + rK$$

subject to

$$y = f(L, K)$$

Q: What is $\mathcal{L}$?

$\mathcal{A}$:

In actuality, the cost function from this problem, $C(w,r,y)$ is a function of not only w and r, but also a given y. You may be interested in finding $\frac{\partial C(w,r,y)}{\partial y}$. Yes, you can take the derivative directly, but sometimes, this result is just very handy. We will talk through one of them.

Here, the Envelope Theorem tells you that:

$$\frac{\partial C(w,r,y)}{\partial y} = \frac{\partial \mathcal{L}}{\partial y} = \lambda^*$$

You will learn eventually about Hotelling's Lemma and Shephard's lemma and their applications to economic theory/results. The Envelope Theorem is what pulls a lot of these together.

Example and exercise. Consider this 'dual' cost minimization problem given a fixed level of utility.

$$\min_{x,y} p_x x + p_y y$$

Subject to

$$u_o = xy$$

Solve for x and y and let's check for the Envelope Theorem holds?