Mathematics for Economists: Lecture 8

Juuso Välimäki

Aalto University School of Business

Spring 2021

This lecture covers

1. Introduction to inequality constrained optimization
2. Karush-Kuhn-Tucker first-order conditions
3. Concave programming
4. Utility maximization problem
5. Cost minimization problem

Introduction: Utility maximization

- Choose a consumption vector $(x_1, ..., x_n) \geq 0$ to:

$$\max_{x_1,\dots,x_n} u(x_1, ..., x_n)$$

in the feasible set given by:

$$\sum_{i=1}^{n} p_i x_i \leq w,$$

$$-x_i \leq 0 \text{ for all } i.$$

- Solution of the problem: demand functions

$$(x_1(p_1, ..., p_n, w), ..., x_n(p_1, ..., p_n, w)).$$

Introduction: Cost minimization

- A firm chooses its inputs k, l to minimize the cost of reaching a production target of $\overline{q}$ at given input prices $r, w > 0$.
- The production function is assumed to be a strictly increasing and quasiconcave function $f(k, l)$.

$$c(r, w; \bar{q}) := \min_{(k,l)} rk + wl$$

subject to

$$\overline{q} \leq f(k, l),$$

$$k \geq 0, \quad l \geq 0.$$

- Solution: conditional factor demands $k(r, w, \bar{q}), l(r, w, \bar{q})$ and cost function $c(r, w; \bar{q})$ giving the minimal cost to achieve production target $\bar{q}$.

Optimization with inequality constraints: Formulation

- The most important class of optimization problems in economics considers maximizing (or minimizing) an objective function $f : \mathbb{R}^n \to \mathbb{R}$ subject to k inequality constraints.
- In these problems, the feasible set takes the form

$$F = \{\boldsymbol{x} \in \mathbb{R}^n : h(\boldsymbol{x}) \leq 0\},$$

where $h(\boldsymbol{x}) \leq 0$ can written more fully as:

$$\begin{aligned} h_1(x_1, ..., x_n) &\leq 0 \\ \vdots \quad & \quad \vdots \\ h_k(x_1, ..., x_n) &\leq 0 \end{aligned}$$

Notice that we can incorporate equality constrains into these problems since $\{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) = 0\}$ is the same set as $\{\boldsymbol{x} \in \mathbb{R}^n : g(x) \leq 0, \quad -g(\boldsymbol{x}) \leq 0\}$.

Towards first-order conditions

- Our first goal is to get the first-order conditions for an optimum.
- We say that an inequality constraint $h_j(x_1, ..., x_n) \leq 0$ is binding at $\hat{\boldsymbol{x}}$ if $h_j(\hat{\boldsymbol{x}}) = 0$.
- If $h_j(\hat{\boldsymbol{x}}) < 0$, then we say that the constraint is not binding.
- A non-binding constraint does not restrict the feasible directions for small changes in $\hat{\boldsymbol{x}}$.
- (As with equality constraints) for binding inequality constraints $h_j(\hat{\boldsymbol{x}})$, the feasible directions $\Delta \boldsymbol{x}$ are given again by:

$$Dh_j(\hat{x})\Delta \boldsymbol{x} \leq 0.$$

Towards first-order conditions

- Non-binding constraints can be ignored. The problem in general is that we do not know a priori which constraints are binding and which are not.
- Let's write the Lagrangean function for the optimization problem as before:

$$\mathcal{L}(\boldsymbol{x}, \lambda_1, ..., \lambda_k) = f(\boldsymbol{x}) - \sum_{j=1}^{k} \lambda_j h_j(\boldsymbol{x}).$$

- I have adopted the notation for the textbook to denote the Lagrange multipliers in inequality constrained problems by λ_j.
- Both binding and non-binding constraints can be handled by the following complementary slackness condition. For all j, we have:

$$\hat{\lambda}_j h_j(\hat{x}_1, ..., \hat{x}_n) = 0.$$

Towards first-order conditions

- This simply says that if $h_j(\hat{\boldsymbol{x}}) < 0$, then $\hat{\lambda}_j = 0$ and the constraint vanishes from the Lagrangean.
- If the constraint binds, then $h_j(\hat{\boldsymbol{x}}) = 0$ and the complementary slackness is also satisfied.
- Based on these considerations, we formulate the first order conditions for $(\hat{x}, \hat{\lambda})$ as follows.
- We consider a point where the constraint qualification holds (i.e. the derivatives of the binding constraints are linearly independent so that we can use implicit function theorem).

Kuhn-Tucker or Karush-Kuhn-Tucker conditions

- The first-order conditions for the problem also known as the Kuhn-Tucker or Karush-Kuhn-Tucker conditions for the problem are given by:

$$\frac{\partial \mathcal{L}}{\partial x_i}(\hat{\boldsymbol{x}}, \hat{\lambda}) = \frac{\partial f}{\partial x_i}(\hat{\boldsymbol{x}}) - \sum_{j=1}^{k} \hat{\lambda}_j \frac{\partial h_j}{\partial x_i}(\hat{\boldsymbol{x}}) = 0 \text{ for all } i,$$

$$\hat{\lambda}_j h_j(\hat{\boldsymbol{x}}) = 0 \text{ for all } j \in \{1, ..., k\},$$

$$\hat{\lambda}_j \geq 0 \text{ for all } j \in \{1, ..., k\},$$

$$h_j(\hat{\boldsymbol{x}}) \leq 0 \text{ for all } j \in \{1, ..., k\}.$$

First-order necessary conditions

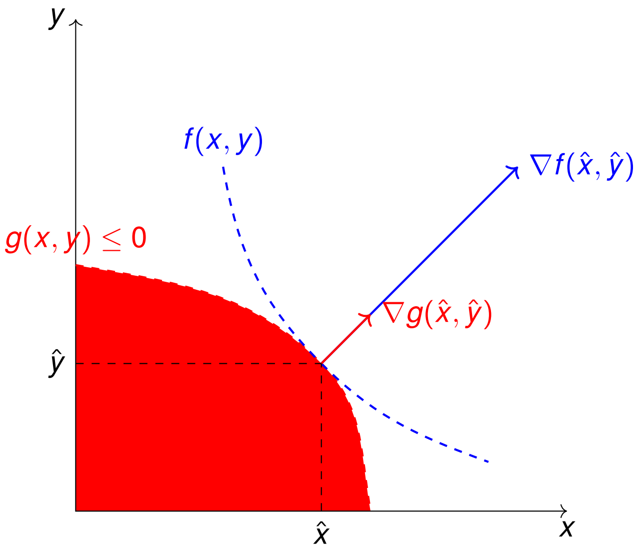
- Let me sum up: at the optimal point $\hat{\boldsymbol{x}}$, we need

  i) the usual first-order condition for the Lagrangean with respect to the choice variables.

  ii) we need that $\hat{\boldsymbol{x}}$ be feasible, i.e. $h_j(\hat{\boldsymbol{x}}) \leq 0$ for all j,

  iii) the complementary slackness conditions, and the non-negativity of the multipliers.

Figure: Single inequality constraint



Optimization with inequality constraints

- We have not discussed the non-negativity of the multipliers yet, but it is easy to see why this must be true in the case of a single inequality constraint.
- Assume constraint qualification, i.e. $Dh(\hat{\boldsymbol{x}}) \neq 0$. By the first order conditions with respect to the x_i, we see that as before,

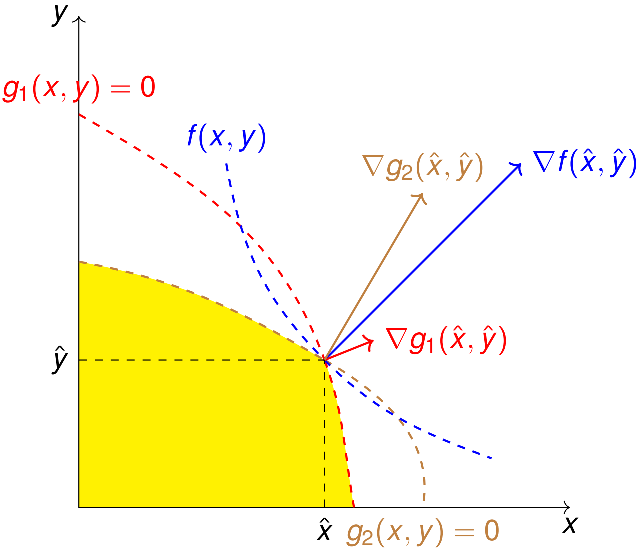
$$\nabla f(\hat{\boldsymbol{x}}) = \lambda \nabla h(\hat{\boldsymbol{x}}).$$

- If the multiplier was strictly negative at an optimal point $\hat{\boldsymbol{x}}$, where the constraint binds, then

$$Dh(\hat{\boldsymbol{x}}) \nabla f(\hat{\boldsymbol{x}}) = \lambda \nabla h(\hat{\boldsymbol{x}}) \cdot \nabla h(\hat{\boldsymbol{x}}) \leq 0.$$

- Hence movement in the direction of the fastest increase of f is feasible and $\hat{\boldsymbol{x}}$ cannot be an optimum unless $\nabla f(\hat{\boldsymbol{x}}) = 0$. But in this case, $\hat{\lambda} = 0$ since $\nabla h(\hat{\boldsymbol{x}}) \neq 0$ by constraint qualification.
- The general case for the positive sign of the multipliers is proved using either separating hyperplane theorem or Farkas' Lemma and it is left for future studies.

Figure: Two inequality constraints



Concave programming

Consider i) maximization problems where the objective function is quasiconcave and ii) minimization problems where the objective function is quasiconvex.

For each of these cases, we assume that the constraint functions h_j are quasiconvex so that the feasible set that is given as the intersection of lower level sets of these functions is convex.

We are now ready to see why the first-order conditions are sufficient for maxima of quasiconcave functions with a non-vanishing derivative on a convex set.

Concave programming

Recall from Lecture 6 that a differentiable function f on a convex set X is quasiconcave if and only if for all $\boldsymbol{x}, \boldsymbol{y} \in X$:

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) \Rightarrow Df(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}) \geq 0.$$

This implies the following (almost converse) result:

Proposition

Suppose $Df(\boldsymbol{x})$ is non-zero for all $\boldsymbol{x} \in X$ and f is quasiconcave on X. Then $\hat{\boldsymbol{x}}$ is a global maximum for f on X if $Df(\hat{\boldsymbol{x}})(\boldsymbol{y} - \hat{\boldsymbol{x}}) \leq 0$ for all $\boldsymbol{y} \in X$

Concave programming

Theorem

Suppose that f is quasiconcave and $Df(\boldsymbol{x}) \neq 0$ on a the convex set $X = \{\boldsymbol{x} \in \mathbb{R}^n : h_j(\boldsymbol{x}) \leq 0 \text{ for } j \in \{1, ..., k\}\}$, where each $h_j(\boldsymbol{x})$ is a quasiconvex function. Then any point satisfying the first-order conditions is a global maximum for f on X.

Concave programming

Proof. Write the first-order condition with respect to x as:

$$Df(\hat{\boldsymbol{x}}) - \sum_{j=1}^{k} \hat{\lambda}_j Dh_j(\hat{\boldsymbol{x}}) = 0. \tag{1}$$

Multiply on the right by $(\boldsymbol{y} - \hat{\boldsymbol{x}})$ to get

$$Df(\hat{\boldsymbol{x}})(\boldsymbol{y} - \hat{\boldsymbol{x}}) - \sum_{j=1}^{k} \hat{\lambda}_j Dh_j(\hat{\boldsymbol{x}})(\boldsymbol{y} - \hat{\boldsymbol{x}}) = 0. \tag{2}$$

Concave programming

Feasible directions for binding constraints satisfy: $Dh_j(\hat{\boldsymbol{x}})(\boldsymbol{y} - \hat{\boldsymbol{x}}) \leq 0$. For nonbinding constraints, $\hat{\lambda}_j = 0$. Therefore since $\hat{\lambda}_j \geq 0$ for all j, we have

$$\hat{\lambda}_j Dh_j(\hat{\boldsymbol{x}})(\boldsymbol{y} - \hat{\boldsymbol{x}}) \leq 0 \text{ for all } j.$$

Thus by equation (2), we see that

$$Df(\hat{\boldsymbol{x}})(\boldsymbol{y} - \hat{\boldsymbol{x}}) \leq 0$$

for all feasible $\boldsymbol{y}$. Therefore by the proposition above, $f(\hat{\boldsymbol{x}}) \geq f(\boldsymbol{y})$ for all feasible $\boldsymbol{y}$.

Utility maximization problem (UMP)

- A consumer allocates her budget of $w > 0$ to n goods.
- Her consumption vector is an element of the positive orthant of the n Euclidean space $X = \{\boldsymbol{x} \in \mathbb{R}^n_+\}$.
- We assume that the consumer has a continuous utility function $u(x)$ defined on X.

$$\boldsymbol{p} \cdot \boldsymbol{x} \leq w \text{ or } \sum_{i=1}^{n} p_i x_i \leq w,$$

where $\boldsymbol{p} = (p_1, ..., p_n) > 0$ is the vector of strictly positive prices for the goods.

Utility maximization problem (UMP)

Maximize

$$u(x_1, ..., x_n)$$

subject to

$$\sum_{i=1}^{n} p_i x_i \leq w,$$

$$x_i \geq 0 \text{ for all } i.$$

Alternatively subject to

$$\sum_{i=1}^{n} p_i x_i - w \leq 0,$$

$$-x_i \leq 0 \text{ for all } i.$$

Utility maximization problem (UMP)

- To see that the feasible set is bounded, let $p^{min} = \min_j p_j$ (i.e. one of the smallest prices p_j).
- Then we know that for all feasible x, we have $p_i x_i \leq w$ for all i since $x_i \geq 0$ and $p_i > 0$ for all i.
- Therefore for all feasible x, $x_i \leq \frac{w}{p^{min}}$ for all i so that the feasible set is bounded since $0 \leq x_i \leq \frac{w}{p^{min}}$ for all i.

Utility maximization problem (UMP)

- To see that the feasible set is closed, we need to show that all limit points of the feasible belong to the feasible set.
- We show this by arguing that when y is not in the feasible set, it is not a limit point.
- If y is not feasible, then either $y_i < 0$ for some i or $\sum_i p_i y_i > w$.
- In both cases all points in a small enough neighborhood of y in infeasible. In the first case, $B^{\varepsilon}(y)$ with $\varepsilon < -\min_i y_i$, in the second, $\varepsilon < \frac{\sum_i p_i y_i - w}{\max_i p_i}$.
- Weiertrass' theorem guarantees that a maximum exists. The solution is called the Marshallian demand (demand as a function of prices and income).

UMP: Lagrangean

- Since the constraint functions are linear, the feasible set is convex.
- If u is strictly increasing (as we usually assume) and quasiconcove, then the first order Kuhn-Tucker conditions are necessary and sufficient for optimum.
- In words, whenever we find a point satisfying the K-T conditions, we have solved the problem.
- Lagrangean:

$$\mathcal{L}(x, \lambda) = u(x) - \lambda_0 \left[\sum_{i=1}^{n} p_i x_i - w \right] + \sum_{i=1}^{n} \lambda_i x_i$$

UMP: K-T conditions

- The first-order K-T conditions are:

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial u(x)}{\partial x_i} - \lambda_0 p_i + \lambda_i = 0 \text{ for all } i, \tag{3}$$

$$\lambda_0 \left[\sum_{i=1}^{n} p_i x_i - w \right] = 0, \tag{4}$$

$$\lambda_i x_i = 0 \text{ for all } i, \tag{5}$$

$$\sum_{i=1}^{n} p_i x_i - w \leq 0, \tag{6}$$

$$-x_i \leq 0 \text{ for all } i, \tag{7}$$

$$\lambda_i \geq 0 \quad i \in \{0, 1, ..., n\}. \tag{8}$$

UMP: Simplifying the K-T conditions

- If the utility function has a strictly positive partial derivative for some x_i at the optimum, then the budget constraint must bind and $\lambda_0 > 0$.
- This follows immediately from the first line of the K-T conditions.
- For the other inequality constraints, consider the partial derivatives at $\boldsymbol{x} \in X$with $x_i \to 0$ for some i.
- If

$$\lim_{x_i \to 0} \frac{\partial u(\boldsymbol{x})}{\partial x_i} = \infty,$$

  then the first line of the K-T conditions implies that at optimum $x_i > 0$.
- If this is true for all i, then we can ignore the non-negativity constraints and we are effectively back to a problem with a single equality constraint.
- If $\frac{\partial u(\boldsymbol{x})}{\partial x_i} < \infty$ for $\boldsymbol{x} = (x_i, \boldsymbol{x}_{-i}) = (0, \boldsymbol{x}_{-i})$, then we must also consider corner solutions where $x_i = 0$ at optimum.

UMP: Interior solutions to K-T conditions

- For interior solutions $x_i > 0$ for all i, we get from the first equation by eliminating λ the familiar condition:

$$\frac{\frac{\partial u(\boldsymbol{x})}{\partial x_i}}{\frac{\partial u(\boldsymbol{x})}{\partial x_k}} = \frac{p_i}{p_k}. \tag{9}$$

- This is of course the familiar requirement that $MRS_{x_i,x_k} = \frac{p_i}{p_k}$ that we saw in Principles of Economics 1.
- Now we see that the same condition extends for many goods and the economic intuition is exactly the same.
- The price ratio gives the marginal rate of transformation between the different goods and at an interior optimum, that rate must coincide with the marginal rate of substitution.

UMP: Interior solutions to K-T conditions

- By multiplying these equations by $p_k \frac{\partial u(\boldsymbol{x})}{\partial x_k}$, we can write the first order conditions for an interior solution as:

$$p_k \frac{\partial u(\boldsymbol{x})}{\partial x_1} - p_1 \frac{\partial u(\boldsymbol{x})}{\partial x_k} = 0 \text{ for all } k, \sum_{i=1}^{n} p_i x_i - w = 0. \tag{10}$$

- In this equation system, we have n endogenous variables $x_1, ..., x_n$ and $n+1$ exogenous variables $p_1, ..., p_n, w$.
- We want to examine the comparative statics of $x(p, w)$, for example $\frac{\partial x_i(\boldsymbol{p},w)}{\partial p_i}$, $\frac{\partial x_i(\boldsymbol{p},w)}{\partial p_j}$ and $\frac{\partial x_i(\boldsymbol{p},w)}{\partial w}$.

UMP: Interior solutions to K-T conditions

- In words, what happens to the demand for one good when its own price changes, when other goods prices change and when income changes.
- Straight application of Implicit function theorem is cumbersome.
- In next week's lectures, we'll do this via duality between UMP and expenditure minimization. Here, tackle easy cases where the optimum can be solved explicitly.

Utility maximization: Cobb-Douglas utility function

- Perhaps the most used functional form in economics is the Cobb-Douglas function
$$u(x) = x^{\alpha} y^{1-\alpha},$$
for some $\alpha \in (0, 1)$.
- The distinguishing feature of this form is that the function is homogenous of degree 1.
- You can check with the Hessian matrix (as an exercise) that $u(x, y)$ is concave and therefore also quasiconcave.

Utility maximization: Cobb-Douglas utility function

- Both marginal utilities are strictly positive at all $(x, y) > (0, 0)$ and

$$\lim_{x \to 0} \frac{\partial u(x, \bar{y})}{\partial x} = \lim_{y \to 0} \frac{\partial u(\bar{x}, y)}{\partial y} = \infty,$$

  for $\bar{x}, \bar{y} > 0$. Since $x = y = \epsilon$ is feasible for small enough ϵ and $u(\epsilon, \epsilon) = \epsilon > 0 = u(0, 0)$, we know that even though $(\hat{x}, \hat{y}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)$ is a critical point, it is not a maximum.
- Any other point satisfying the KT- conditions has $Df(x, y) \neq 0$ and therefore will be the optimum.

UMP: Cobb-Douglas utility function

- First-order conditions w.r.t. x, y imply that any solution (different from $(0,0)$) is interior and the budget constraint binds.
- The requirement that $MRS_{x,y} = \frac{p_x}{p_y}$ gives:

$$\frac{\alpha y}{(1-\alpha)x} = \frac{p_x}{p_y} \text{ or } p_x x = \frac{1-\alpha}{\alpha} p_y y.$$

- From budget constraint:

$$p_x x + p_y y = w,$$

- Substituting, we get:

$$x(p_x, p_y, w) = \frac{\alpha w}{p_x}, \text{ and } \qquad y(p_x, p_y, w) = \frac{(1-\alpha)w}{p_y}.$$

UMP: Cobb-Douglas utility function

- For the Cobb-Douglas utility function, you get the result that the expenditure shares $\frac{p_x x}{w} = \alpha$ and $\frac{p_y y}{w} = 1 - \alpha$ do not depend on prices or w.
- This extends easily to the case with n goods and $u(x) = x_1^{\alpha_1} \dots x_n^{\alpha_n}$ with $\alpha_i > 0, \sum_i \alpha_i = 1$ at prices $p = (p_1, ..., p_n)$. Then you have:

$$x_i(p, w) = \frac{\alpha_i w}{p_i}.$$

- This is not very realistic.
- The rich and the poor use their budgets very differently.

UMP: Stone-Geary utility function

- One way to get more realistic consumption patters is to define the utility function for consumptions above a level needed for subsistence.
- Let $\underline{\boldsymbol{x}} = (\underline{x_1}, ..., \underline{x_n})$ be the levels of each good needed for survival and assume that $w \geq \boldsymbol{p} \cdot \underline{\boldsymbol{x}}$.
- The utility function for $\boldsymbol{x} \in \mathbb{R}^n$ such that $x_i \geq \underline{x_i}$ is of Cobb-Douglas -like form:

$$u(\boldsymbol{x}) = (x_1 - \underline{x_1})^{\alpha_1} \ldots (x_n - \underline{x_n})^{\alpha_n},$$

  where $0 < \alpha_i < 1$ for all i and $\sum_{i=1}^{n} \alpha_i = 1$.
- Notice that the marginal utility for good i is infinite if $x_i = \underline{x_i}$ and that the utility function is strictly increasing in all of its components.
- Hence we still have an interior solution and the budget constraint binds.

UMP: Stone-Geary utility function

- We get as above:

$$\frac{\frac{\partial u(\boldsymbol{x})}{\partial x_i}}{\frac{\partial u(\boldsymbol{x})}{\partial x_k}} = \frac{\alpha_i(x_k - \underline{x_k})}{\alpha_k(x_i - \underline{x_i})} = \frac{p_i}{p_k} \text{ for all } i, k,$$

$$\sum_{i=1}^{n} p_i x_i = w.$$

- Taking $k = 1$, we get that

$$x_i - \underline{x_i} = \frac{\alpha_i p_1}{\alpha_1 p_i}(x_1 - \underline{x_1}) \text{ for all } i. \tag{11}$$

- Multiplying both sides by p_i and summing over i gives:

$$\sum_{i=1}^{n} p_i(x_i - \underline{x_i}) = \frac{p_1 \sum_{i=1}^{n} \alpha_i}{\alpha_1}(x_1 - \underline{x_1}).$$

UMP: Stone-Geary utility function

- So we can solve:

$$x_1 - \underline{x_1} = \frac{\alpha_1 (w - \sum_{i=1}^n p_i \underline{x_i})}{p_1},$$

  where we used the budget constraint $\sum_{i=1}^n p_i x_i = w$ and $\sum_{i=1}^n \alpha_i = 1$
- By (11), we see that

$$x_i - \underline{x_i} = \frac{\alpha_i (w - \sum_{j=1}^n p_j \underline{x_j})}{p_i}.$$

- The consumer uses a constant fraction of her excess income (above what is needed for the necessities $\underline{x}$) in constant shares given by the α_i.
- Since the poor have less excess wealth, their consumption fractions are closer to the ones given by the subsistence levels $\beta_i := \frac{\underline{x_i}}{\sum_i \underline{x_i}}$.