Advanced Micro: Recitation 5
Homogeneity and Kuhn-Tucker

Motaz Al-Chanati

March 1, 2019

1 Homogeneous Functions

1.1 Definition

Homogeneity For any scalar k, a real-valued function $f(x_1, ..., x_n)$ is homogeneous of degree (h.o.d.) k if:

$$f(\lambda x_1, ..., \lambda x_n) = \lambda^k f(x_1, ..., x_n), \forall \lambda > 0$$

Cone A cone is a set C such that if $x \in C$, then $\lambda x \in C, \forall \lambda > 0$.

We will always assume that the domain of a homogeneous function must be a cone (so that we can actually evaluate λx) We have already seen this idea in class a number of times:

- Marshallian demand is h.o.d. 0 in prices and income
- Hicksian demand is h.o.d. 0 in prices
- The expenditure function is h.o.d. 1 in prices
- Cobb-Douglas function of the form $f(x) = \prod_{i=1}^n x_i^{\alpha_i}$ is h.o.d. $\sum_{i=1}^n \alpha_i$

1.2 Properties

Let's go through some useful properties of homogeneous functions

Theorem 1. *Let $f(x)$ be a C^1 function on an open cone in $\mathbb{R}^n$. If f is h.o.d. k, its first order partial derivatives are h.o.d. $k-1$*

Proof. Start with the fact that f is h.o.d. k and then take the partial derivative with respect to x_i to both sides (using the chain rule on the LHS)

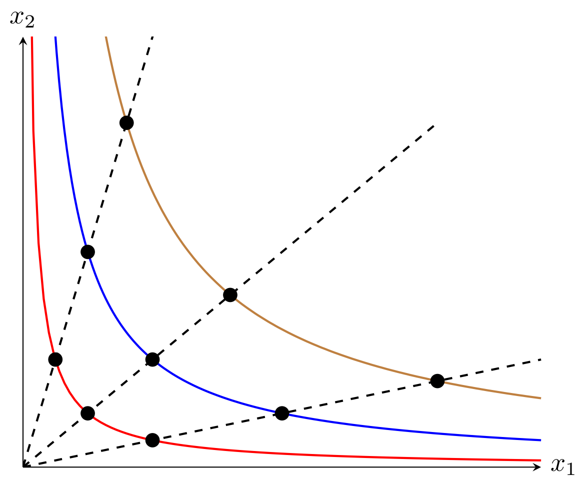
$$\begin{aligned} f\left(\lambda x_1, ..., \lambda x_i, ..., \lambda x_n\right) &= \lambda^k f\left(x_1, ..., x_i, ..., x_n\right) \\ \frac{\partial f}{\partial x_i}\left(\lambda x_1, ..., \lambda x_i, ..., \lambda x_n\right) \cdot \frac{\partial(\lambda x_i)}{\partial x_i} &= \frac{\partial}{\partial x_i} \lambda^k f\left(x_1, ..., x_i, ..., x_n\right) \\ \frac{\partial f}{\partial x_i}\left(\lambda x_1, ..., \lambda x_i, ..., \lambda x_n\right) \cdot \lambda &= \lambda^k \frac{\partial f}{\partial x_i}\left(x_1, ..., x_i, ..., x_n\right) \end{aligned}$$

$$\therefore \frac{\partial f}{\partial x_i}(\lambda x) = \lambda^{k-1}\frac{\partial f}{\partial x_i}(x)$$

□

Consider a function $q = f(x)$ of h.o.d. k. Consider the level curve at some point $\tilde{q}$. Then we can find the level curve at $\lambda^k \tilde{q}$ by simply scaling up the points from the $\tilde{q}$ level curve by λ. We can easily see this since for any $\tilde{x}$ such that $\tilde{q} = f(\tilde{x})$, we have $f(\lambda\tilde{x}) = \lambda^k f(\tilde{x}) = \lambda^k \tilde{q}$. Geometrically, this is translating the points on the $\tilde{q}$ level curve by a factor of k along rays from the origin. You can see this in Figure 1.

Figure 1: Level Curves and Rays from the Origin



A consequence of this is the following result:

Theorem 2. *Let $q = f(x)$ be a C^1 homogeneous function in the first quadrant in $\mathbb{R}^2$. The tangent planes to the level curves of f have constant slope along each ray from the origin.*

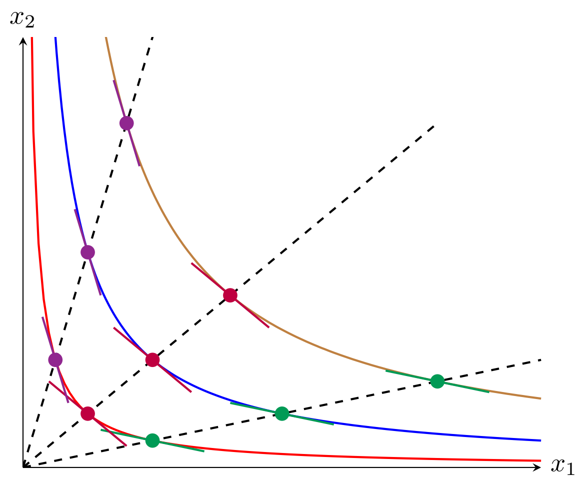
Proof. Take two bundles (x_1^0, x_2^0) and $(x_1^1, x_2^1) = \lambda(x_1^0, x_2^0)$. These are two points that lie on the same ray from the origin. Let $f_1 = \frac{\partial f}{\partial x_1}$ and $f_2 = \frac{\partial f}{\partial x_2}$. By the Implicit Function Theorem, the slope of the level curve at (x_1^1, x_2^1) is:

$$\begin{aligned}
x_2'(x_1)|_{(x_1^1,x_2^1)} &= -\frac{f_1(x_1^1, x_2^1)}{f_2(x_1^1, x_2^1)} && \text{(by IFT)} \\
&= -\frac{f_1(\lambda x_1^0, \lambda x_2^0)}{f_2(\lambda x_1^0, \lambda x_2^0)} && \text{(definition of } x^1 \text{ bundle)} \\
&= -\frac{\lambda^{k-1} f_1(x_1^0, x_2^0)}{\lambda^{k-1} f_2(x_1^0, x_2^0)} && \text{(by above theorem)} \\
&= -\frac{f_1(x_1^0, x_2^0)}{f_2(x_1^0, x_2^0)} \\
&= x_2'(x_1)|_{(x_1^0,x_2^0)} && \text{(by IFT, same slope as } x^0 \text{ bundle)}
\end{aligned}$$

□

We can visually capture this idea in Figure 2.

Figure 2: Slopes of Level Curves along Rays from the Origin



It is useful to have a calculus condition that is necessary and sufficient for a function to be homogeneous. The necessary condition is known as Euler's theorem: (the proof for the sufficient condition is in the Appendix of Chapter 24 in Simon and Blume)

Theorem 3. ***(Euler's Theorem)*** *Let $f(x)$ be a C^1 function of h.o.d. k on $\mathbb{R}^n$. Then, for all x:*

$$\sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i}(x) = kf(x)$$

Or equivalently, in gradient notation:

$$x \cdot \nabla f(x) = kf(x)$$

Proof. Differentiate with respect to λ and then evaluate

$$\begin{aligned} f(\lambda x) &= \lambda^k f(x) \\ \frac{\partial}{\partial \lambda} f(\lambda x) &= \frac{\partial}{\partial \lambda} \lambda^k f(x) \\ \frac{\partial f}{\partial x_1}(\lambda x)x_1 + ... + \frac{\partial f}{\partial x_n}(\lambda x)x_n &= k\lambda^{k-1} f(x) \end{aligned}$$

Note that this is true for every λ. In particular, we can evaluate this at $\lambda = 1$:

$$\begin{aligned} \frac{\partial f}{\partial x_1}(x)x_1 + ... + \frac{\partial f}{\partial x_n}(x)x_n &= k(1)^{k-1} f(x) \\ x \cdot \nabla f(x) &= kf(x) \end{aligned}$$

$\square$

1.3 Application to Utility Functions

Theorem 2 has a clear application to utility functions - it tells us that for homogeneous utility functions, the MRS is constant along rays from the origin. This has an important implication in the consumer's problem. At an interior solution we set MRS equal to the price ratio, and the consumer chooses the bundle $x(p, y)$.

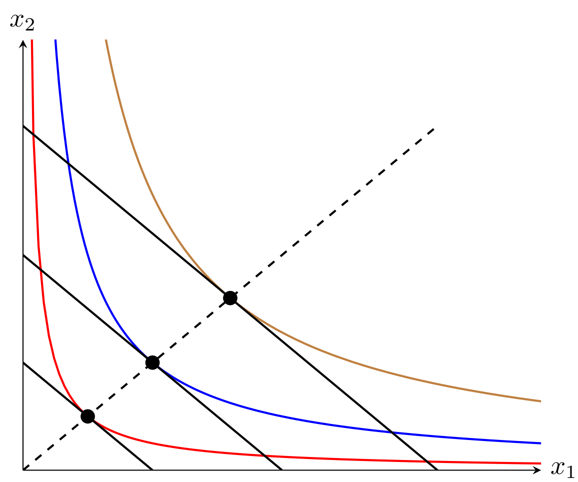
If income increases by a factor of λ, then the new optimal bundle $x(p, \lambda y)$ will have to continue to have the same MRS as before (since the prices have stayed the same). Therefore, the new optimal bundle is where the budget line $p \cdot x = \lambda y$ intersects with the ray from the origin passing through the original optimal bundle $x(p, y)$. This would work for any income level, thus showing us that the income expansion path for a homogeneous utility function is indeed that ray from the origin.

Moreover, since the budget line has had a parallel shift, it has to also be the case that $x(p, \lambda y) = \lambda x(p, y)$. To see this, note that the ray from the origin through the original bundle is represented by the line $x_2 = \frac{x_2(p,y)}{x_1(p,y)} x_1$. Plug this into the new budget constraint to see where they intersect:

$$\begin{aligned} p_1 x_1 + p_2 x_2 &= \lambda y \\ p_1 x_1 + p_2 \frac{x_2(p, y)}{x_1(p, y)} x_1 &= \lambda y \\ \left(\frac{p_1 x_1(p, y) + p_2 x_2(p, y)}{x_1(p, y)} \right) x_1 &= \lambda y \\ \left(\frac{y}{x_1(p, y)} \right) x_1 &= \lambda y \\ \therefore x_1 &= \lambda x_1(p, y) \end{aligned}$$

And similarly for x_2. Therefore, for a homogeneous utility function (of any degree), the demand will be h.o.d. 1 in income. All these ideas are captured in the diagram below.

Figure 3: Income Changes with Homogeneous Utility Function



Since we know this, we can draw one final conclusion regarding the income elasticity of demand. For a fixed price p, the above shows us that we can express demand as a linear function of income: $x_i(p, y) = a_i y, \forall i$. Therefore, the elasticity is:

$$\eta_i = \frac{\partial x_i(p, y)}{\partial y} \cdot \frac{y}{x_i} = a_i \cdot \frac{y}{a_i y} = 1$$

Therefore, for a homogeneous utility function, the income elasticity of demand is a constantly unit-elastic (i.e. if income increases by 1%, then quantity demanded increases by 1%).

We can summarize these ideas as follows:

Theorem 4. *For a utility function $u(x)$ on $\mathbb{R}^n$ that is h.o.d. k, then:*

1. *The MRS is constant along rays from the origin*
2. *The income expansion path are rays from the origin*
3. *The corresponding demand depends linearly on income*
4. *The income elasticity of demand is identically one*

Finally, recall that in class we have seen Engel and Cournot aggregation.

Engel:

$$\sum_{i=1}^{n} s_i \eta_i = 1$$

Cournot:

$$\sum_{i=1}^{n} s_i \varepsilon_{ij} = -s_j$$

We can have get another result using Euler's theorem. Since we know the Marshallian demand for any good i, $x_i(p, y)$, is h.o.d. 0 in prices and income:

$$\left(\sum_{j=1}^{n} p_j \frac{\partial x_i(p,y)}{\partial p_j}\right) + y \frac{\partial x_i(p,y)}{\partial y} = 0 \cdot x_i(p,y) \qquad \text{(Euler's theorem)}$$

$$\left(\sum_{j=1}^{n} \frac{p_j}{x_i} \frac{\partial x_i(p,y)}{\partial p_j}\right) + \frac{y}{x_i} \frac{\partial x_i(p,y)}{\partial y} = 0 \qquad \text{(divide by } x_i\text{)}$$

$$\sum_{j=1}^{n} \varepsilon_{ij} + \eta_i = 0 \qquad \text{(elasticity definitions)}$$

This tells that if all prices and income change by the same percentage, then the quantity demanded will not change - which is exactly what h.o.d. 0 in prices and income tells us too.

2 Kuhn-Tucker Conditions[1]

In most of our optimization problems this semester, the constraint set is actually given by a set of inequality constraints. For example, in its most general form, the consumer's problem is:

$$\max_x u(x) \text{ s.t. } px \leq y; x \geq 0$$

Let me say two things about approaching problems with inequality constraints:

- Frequently we can argue some constraints must bind. For example, if utility is strictly increasing, then we must have $px = y$. Thus we can move forward as though we had an equality constraint
- Frequently we will ignore some constraints. For example, suppose we ignore the positivity constraints and solve the problem $\max_x u(x)$ s.t. $px = y$. If the resulting solution x^* satisfies $x^* \geq 0$, we've found the solution to the original problem. Why? Because x^* is optimal over a set that contains the feasible set of x's, and so must be optimal over the feasible set itself. BUT, if we found an x^* that violated positivity, we would need to go back and impose the positivity constraints.

[1] This section is taken from David Thompson's 2018 recitation notes

It's often easier to argue which constraints will bind and won't, and then use Lagrange's method on the resulting equality-constrained problem. But sometimes we can't do this, in which case we turn to the Kuhn-Tucker conditions (Jehle and Reny Theorem A2.20).

Theorem 5. *Let $f(x)$ and $g^j(x)$ for $j = 1, ..., m$ be C^1, real-valued functions defined over some domain $D \subset \mathbb{R}^n$. Let x^* be an interior point of D and suppose that x^* solves*

$$\max_{x \in D} f(x) \text{ s.t. } g_j(x) \leq 0, j = 1, ..., m$$

If the gradient vectors $\nabla g_j(x^*)$ associated with constraints j that bind at x^* are linearly independent, then there is a unique vector $\lambda^* \in \mathbb{R}^m$, such that (x^*, λ^*) satisfy the Kuhn-Tucker conditions:

$$\begin{aligned} \frac{\partial f(x^*)}{\partial x_i} - \sum_{j=1}^{m} \lambda_j^* \frac{\partial g^j(x^*)}{\partial x_i} &= 0, i = 1, ..., n \text{ (tangency)} \\ g_j(x^*) &\leq 0, j = 1, ..., m \text{ (feasibility)} \\ \lambda_j^* &\geq 0, j = 1, ..., m \text{ (positivity)} \\ \lambda_j^* g^j(x^*) &= 0, j = 1, ..., m \text{ (slackness)} \end{aligned}$$

A few comments on these conditions

- All constraints are stated as $\leq$ constraints. If you're given $\geq$ constraints, multiply by -1. (Different authors use different conventions)
- The first two conditions are exactly what we had with Lagrange's method. Construct the Lagrangian $\mathcal{L}(x, \lambda) = f(x) - \lambda g(x)$ and look for points that satisfy the FOC
- Positivity has a nice interpretation. From the envelope theorem we know Lagrange multipliers measure the change in the value function when a constraint is relaxed by one unit. Positivity simply says the value function doesn't decrease as the constraints are relaxed
- Slackness also has a nice interpretation. If constraint j is binding, $g_j(x^*) = 0$, so slackness holds. If not, then slackness tells us $\lambda_j^* = 0$. This says that if a constraint doesn't bind, relaxing the constraint a little bit will not affect the value function (since x^* isn't "close" to the constraint)

The KT conditions are *necessary* conditions, but not sufficient. The general approach to these problems is to find all points which satisfy the KT conditions, and then find which one of those points generates the largest value of the function.[2] Alternatively, if we have sufficient concavity on our objective and constraint functions, the KT conditions become sufficient as well.

2.1 KT Example

Consider the following problem:

$$\min_{x,y} (x-4)^2 + (y-4)^2 \text{ s.t. } \begin{cases} x + y \leq 4 \\ x + 3y \leq 9 \end{cases}$$

[2] There's also the annoyance of the constraint qualification, which we won't worry about in this class. If you're curious, read Simon and Blume 18.2. If the constraint is linear (as it usually is for us), then the constraint qualification will be satisfied

Notice the KT conditions were stated in terms of a maximum. However, the (x, y) that solves this problem will also solve the associated problem:

$$\max_{x,y} -(x-4)^2 - (y-4)^2 \text{ s.t. } \begin{cases} x + y \leq 4 \\ x + 3y \leq 9 \end{cases}$$

We'll proceed as follows:

- Construct the Lagrangian: $\mathcal{L}(x, y, \lambda) = -(x-4)^2 - (y-4)^2 - \lambda_1(x + y - 4) - \lambda_2(x + 3y - 9)$
- Write down the KT conditions (there are a lot!):

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x} = -2(x-4) - \lambda_1 - \lambda_2 &= 0 && (1)\\
\frac{\partial \mathcal{L}}{\partial y} = -2(y-4) - \lambda_1 - 3\lambda_2 &= 0 && (2)\\
\lambda_1 &\geq 0 && (3)\\
\lambda_2 &\geq 0 && (4)\\
\lambda_1(x + y - 4) &= 0 && (5)\\
\lambda_2(x + 3y - 4) &= 0 && (6)\\
x + y &\leq 4 && (7)\\
x + 3y &\leq 9 && (8)
\end{aligned}$$

- Now go through cases to look for points that satisfy the KT conditions
- Case 1: Neither constraint binds.
  - In this case, slackness (eqs 3 and 4) tells us $\lambda_1 = \lambda_2 = 0$
  - Given this, the FOC (eqs 1 and 2) imply $x = y = 4$
  - However, this violates the constraints: $x + y = 8 > 4$
  - No point satisfies the KT conditions in this case.
- Case 1: Constraint 1 binds, Constraint 2 doesn't
  - Slackness implies $\lambda_2 = 0$
  - Combining eqs 1 and 2 gives $x = y$
  - Using the constraint $x + y = 4$ implies $x = y = 2$
  - Using the FOC we see $\lambda_1 = 4 \geq 0$
  - The second constraint is satisfied: $2 + 3 * 2 = 8 \leq 9$
  - Therefore $(x^*, y^*, \lambda_1^*, \lambda_2^*) = (2, 2, 4, 0)$ satisfy all the KT conditions, and are a candidate for the maximum. We see $f(2, 2) = -8$.
- Case 3: Constraint 2 binds, Constraint 1 doesn't
  - Slackness implies $\lambda_1 = 0$
  - The FOC imply $6(x - 4) = 2(y - 4)$.
  - Combining the above with the constraint $x + 3y = 9$ we see $x = 3.3, y = 1.9$

- However, this implies $x + y = 5.2 > 4$, so constraint 1 is violated. No point satisfies the KT conditions in this case

- Case 4: Both constraints bind
  - In this case, there is only one point that satisfies both constraints: $x = 1.5, y = 2.5$. We see $f(1.5, 2.5) = -8.5 < f(2.2)$, so this point cannot be the optimal point
  - Alternatively, combining the values of x, y with the FOC implies $\lambda_2 = -1$, which violates positivity

Thus there is only one point that satisfies the KT conditions: $(x, y) = (2, 2)$, which is indeed the maximum of this problem subject to the given inequality constraints.