Recitation Notes on Kuhn-Tucker Theorem

Andrea Ciccarone*

November 2022

1 Optimization with Inequality Constraints

1.1 The General Problem

Let us introduce the general problem of maximizing a function under inequality and equality constraints. More specifically, we are interested in the following problem:

$$\max_{x \in X} f(x) \text{ s.t. } g(x) \leq b \text{ and } h(x) = c$$

Where X is an open set in $\mathbb{R}^n$, $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^k$, and $h : X \to \mathbb{R}^m$ are continuous and differentiable functions. These functions are multivariate functions that represent our constraints. More specifically, $h(x) = c$ explicitly rewrites as $h_i(x) = c_i$ for all i :

$$\begin{cases} h_1(x) & = c_1 \\ \vdots & \\ h_m(x) & = c_m \end{cases}$$

In practice, to solve this kind of problems, we often define the **Lagrangian function** of the maximization problem as

$$\begin{aligned} \mathcal{L}(x, \lambda, \mu) &= f(x) - \lambda (g(x) - b) - \mu (h(x) - c) \\ &= f(x) - \sum_{j=1}^{k} \lambda_j (g_j(x) - b_j) - \sum_{l=1}^{m} \mu_l (h_l(x) - c_j) \end{aligned}$$

and λ_j's and μ_l's are called the **Lagrangian multipliers**. We take FOCs wrt to (x, λ, μ) and we check a couple of other conditions that are formalized in the Kuhn-Tucker theorem.

*These notes are partially based on recitation notes by Vinayak Iyer and Cesar Barilla.

1.2 Constraint Qualification

Before we actually state KKT, we need to first introduce **Constraint Qualification**.

Definition 1 *Let X be an open set in $\mathbb{R}^n$, and let $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^k$, and $h : X \to \mathbb{R}^m$ be C^1 functions. Consider the problem*

$$\max_{x \in X} f(x) \;\; \text{s.t. } g(x) \leq b \text{ and } h(x) = c$$

*For a feasible point $\hat{x} \in X$, the inequality constraint $g_j(x) \leq b_j$ is said to be **binding at** $\hat{x}$ iff $g_j(\hat{x}) = b_j$.*

*We say that the **constraint qualification (CQ) holds at** $\hat{x}$ iff the derivatives of all binding constraints*

$$\{\nabla g_j(\hat{x})\}_{\{j: g_j \text{ binding at } \hat{x}\}} \cup \{\nabla h_l(\hat{x})\}_{l=1}^{m}$$

*in $\mathbb{R}^n$ are linearly independent; otherwise we say that the **constraint qualification (CQ) fails at** $\hat{x}$.*

We are going to understand better what CQ means when we introduce KKT. For now, we need to understand that in order to check whether CQ holds at $\hat{x}$ we need to **derive the gradients of the binding constraints** at $\hat{x}$ and then check that they are linearly independent. If that is the case, CQ holds. We are now ready to state KKT.

1.3 Kuhn-Tucker Theorem

Theorem 1 (Kuhn-Tucker) *Let X be an open set in $\mathbb{R}^n$, and let $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^k$, and $h : X \to \mathbb{R}^m$ be C^1 (continuous and differentiable) functions. Consider the problem*

$$\max_{x \in X} f(x) \;\; \text{s.t. } g(x) \leq b \text{ and } h(x) = c$$

If x^* is a maximizer of the problem above, and CQ holds at x^*, then there exists a unique $(\lambda, \mu) \in \mathbb{R}^k_+ \times \mathbb{R}^m$ s.t. the following two conditions hold:

*(1) **First order condition (FOC)**:*

$$\underbrace{\nabla f(x^*)}_{n\times 1} - \underbrace{g'(x^*)}_{n\times k} \cdot \underbrace{\lambda}_{k\times 1} - \underbrace{h'(x^*)}_{n\times m} \cdot \underbrace{\mu}_{m\times 1} = \underbrace{0}_{n\times 1} \;{}^*$$

*(2) **Complementary slackness condition (CSC)**:*

$$h_l(x^*) = c_l$$

for each $l \in \{1, \ldots, m\}$.

$$\lambda_j \geq 0, \text{ with } \lambda_j g_j(x^*) = b_j$$

for each $j \in \{1, \ldots, k\}$.

*Note that this is a set of n equations - one for each x_i.

Notes :

1. The constraints $g(x) \leq b$ are actually a set of k inequalities i.e.

$$\begin{aligned} & g(x) \leq b \\ \iff & \begin{pmatrix} g_1(x) \\ \vdots \\ g_k(x) \end{pmatrix} \leq \begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix} \end{aligned}$$

   and the constraints $h(x) = c$ are actually a set of m equalities i.e.

$$\begin{aligned} & h(x) = c \\ \iff & \begin{pmatrix} h_1(x) \\ \vdots \\ h_m(x) \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \end{aligned}$$

2. The FOC is essentially for each $i \in \{1, \dots, n\}$:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i}(x^*, \lambda, \mu) = & \frac{\partial}{\partial x_i}\left(f(x) - \lambda^T(g(x) - b) - \mu^T(h(x) - c)\right)\Big|_{x^*} = 0 \\ & \frac{\partial f}{\partial x_i}(x^*) - \sum_{j=1}^{k} \lambda_j \frac{\partial g_j}{\partial x_i}(x^*) - \sum_{l=1}^{m} \mu_l \frac{\partial h_l}{\partial x_i}(x^*) = 0 \end{aligned}$$

3. Note that the term $\underbrace{g'(x^*)}_{n \times k} \cdot \underbrace{\lambda}_{k \times 1}$ is nothing but the expression :

$$\underbrace{g'(x^*)}_{n \times k} \cdot \underbrace{\lambda}_{k \times 1} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_k}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix}$$

   and the term $\underbrace{h'(x^*)}_{n \times m} \cdot \underbrace{\mu}_{m \times 1}$ is given by :

$$\underbrace{h'(x^*)}_{n \times m} \cdot \underbrace{\mu}_{m \times 1} = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial x_n} & \cdots & \frac{\partial h_m}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix}$$

4. Simply, the Kuhn-Tucker Theorem states that **if x^* is a maximizer and satisfies the CQ then there exists λ and μ such that (x^*, λ, μ) satisfies FOC+CSC**.

5. **What is the interpretation of the Complementary Slackness Condition?** Remember that the multiplier λ_j is the marginal increase in the maximized objective function due to a slight relaxation of the constraint. The CSC simply says that if a constraint is not binding at the optimum, then if we relax the constraint a little, we should not be increasing the value of the objective function i.e. $\lambda_j = 0$. In a utility

maximizing framework, if the Budget constraint does not bind at the optimum i.e. $p \cdot x^* < m$, then we are leaving some money on the table already. Now if we increase m, then our choice x^* should not change which implies that $U(x^*)$ is unchanged which in turn implies that the multiplier associated with the budget constraint is zero i.e. $\lambda = 0$.

6. In practice one writes down what are generally called the **KKT conditions**:
   - $\frac{\partial f}{\partial x_i}(x^*) - \sum_{j=1}^{k} \lambda_j \frac{\partial g_j}{\partial x_i}(x^*) - \sum_{l=1}^{m} \mu_l \frac{\partial h_j}{\partial x_i}(x^*) = 0$, $\forall i = 1, \ldots, n$
   - $h_l(x) = c_l$ $\forall l = 1, \ldots, m$
   - $\lambda_j \geq 0, g_j(x) \leq b_j$ and $\lambda_j (g_j(x) - b_j) = 0$, $\forall j = 1, \ldots, k$

Some Additional Comments Regarding Implications of the Kuhn-Tucker Theorem

1. It is **ONLY** a necessary condition and **NOT** a sufficient condition.
2. The Kuhn-Tucker Theorem states that **if x^* is a maximizer and satisfies the CQ, then (x^*, λ, μ) must solve FOC+CSC**. Hence we can find x^* by solving for all solutions to the KKT conditions. It does **NOT** however mean that if some x solves **FOC+CSC** then it is a maximizer.
3. **The Kuhn-Tucker Theorem only works at x^* at which the CQ holds**. If the CQ fails at x^*, then there may not exist (x^*, λ, μ) which satisfy the FOC and CSC even if x^* is a maximizer of the problem. So there may exist other feasible points at which the CQ fails which may also be maximizers.
4. When we are finding the solution to an optimization problem, we should find the solutions where the KKT conditions are satisfied. Call these **Type 1** solutions. We should also find all points at which the CQ fails. Call these **Type 2** solutions. Our solution is then found by comparing all these points i.e., **Type 1 + Type 2**.
5. For **sufficiency**, if a point x^* satisfies the KKT conditions and the Lagrangian is concave, then x^* is a maximizer.

Meaning of CQ Note that the FOC implies that:

$$\underbrace{\nabla f(x^*)}_{n\times 1} = \underbrace{g'(x^*)}_{n\times k} \cdot \underbrace{\lambda}_{k\times 1} + \underbrace{k'(x^*)}_{n\times m} \cdot \underbrace{\mu}_{m\times 1}$$

that is, the gradient of f can be written as a linear combination of the gradients of the constraints.

First note that, by CSC, $\lambda_j = 0$ for those constraints which are not binding. More so, since the theorem states that a **unique** (λ, μ) exists, then it must be that the gradient vector of constraints which bind must be Linearly Independent. This is because we know from Linear Algebra that vector is expressed as a linear combination of some linearly independent vectors, then the representation unique.

2 Applications of KKT

Example 1

Consider the following problem :

$$\begin{aligned} &\max\, xy + x^2 \\ &\text{s.t } g_1(x,y) = x^2 + y \leq 2 \\ &\quad\; g_2(x,y) = -y \leq -1 \end{aligned}$$

We will work through how to do it in a systematic procedure. The first thing to note is that a solution exists by Weierstrass' Theorem since the objective function is continuous and the constraint set is closed and bounded. Moreover, first note that the Lagrangian of this problem is :

$$\mathcal{L} = xy + x^2 - \lambda_1\left(x^2 + y - 2\right) - \lambda_2(-y+1)$$

The KT conditions are then:

$$\begin{cases} \mathcal{L}_x = y + 2x - 2\lambda_1 x = 0 \\ \mathcal{L}_y = x - \lambda_1 + \lambda_2 = 0 \\ \lambda_1 \geq 0, \lambda_2 \geq 0 \\ \lambda_1\left(x^2 + y - 2\right) = 0 \\ \lambda_2\left(-y+1\right) = 0 \end{cases}$$

Now we need to consider all the possible cases that we can have:

Case 1 - Both constraints are binding: When both constraints are binding we have $x^2 + y = 2$ and $y = 1$. This implies $x = \pm 1$. Let us first consider $(1,1)$: from the FOCs we have that $\lambda_1 = \frac{3}{2}$ and $\lambda_2 = \frac{1}{2}$. Our first Type 1 candidate is thus $(x,y) = (1,1)$ with $(\lambda_1, \lambda_1) = (\frac{3}{2}, \frac{1}{2})$.

Now we can consider $(-1,1)$. From the FOCs, we can derive $(\lambda_1, \lambda_1) = (\frac{1}{2}, \frac{3}{2})$. This is our second Type 1 candidate. Notice that we still have to check for CQ, we will do that later.

Case 2 - Constraint 1 is binding and Constraint 2 is not binding: Now we have that $x^2 + y = 2$ and $y > 1$. From the CSC we have $\lambda_2 = 0$ so that $x = \lambda_1$. We can plug this in the first FOC and recalling that $y = 2 - x^2$ from the constraint, we get:

$$3x^2 - 2x - 2 = 0$$

Solving this equation yields $x = \frac{1}{3}\left(1 \pm \sqrt{7}\right)$. But since $x = \lambda_1 \geq 0$ it must be $x = \frac{1}{3}\left(1 + \sqrt{7}\right)$. This implies $y = 2 - x^2 = \frac{2}{9}\left(5 - \sqrt{7}\right) < 1$ which violates $y > 1$. Therefore, there are not solution candidates in this case.

Case 3 - Constraint 1 is not binding and Constraint 1 is binding: We now have $y = 1$ and $x^2 + y < 2$. We thus have $\lambda_1 = 0$ and thus from the FOCs $x = -\frac{1}{2}$ which in turn yields $\lambda_2 = \frac{1}{2}$. Our candidate in this case is thus $(x, y) = \left(-\frac{1}{2}, 1\right)$ with $(\lambda_1, \lambda_1) = \left(0, \frac{1}{2}\right)$.

Case 4 - Both constraints are not binding: In this case from the CSC it must be $\lambda_1 = \lambda_2 = 0$ which implies, from the FOCs, that $x = 0$ and thus $y = 0$, contradicting $y > 1$.

Summing up, we have 3 Type 1 candidates. Our favorite candidate is going to be $(1, 1)$ since $f(1, 1) = 2$ which is greater than both $f(-1, 1)$ and $f(-\frac{1}{2}, 1)$.

The only thing we have left is to check points at which CQ may fail. Notice that:

$$\nabla g_1(x, y) = \begin{pmatrix} 2x \\ 1 \end{pmatrix}$$
$$\nabla g_2(x, y) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Notice that since none of these vectors is the 0 vector, the only problem we may have is when both constraints are binding and the vectors are linearly dependent. This occurs if and only if $x = 0$. When constraints are binding x can not be equal to 0 so we can ignore this points. Anyways, $f(x, y) = 0$ when $x = 0$ so it still couldn't be our maximizer. Thus, the unique maximizer of our problem is $(x^*, y^*) = (1, 1)$.

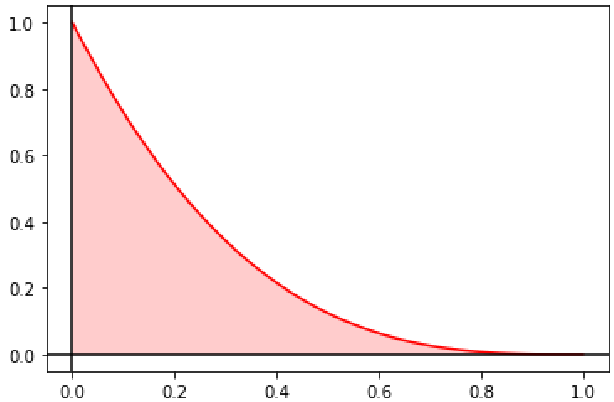
Example 2 (When CQ Really Matters)

Consider the following problem:

$$\begin{aligned} &\max_{x,y} x \\ &\text{s.t } y - (1 - x)^3 \leq 0 \\ &\quad x, y \geq 0 \end{aligned}$$

The feasible set in this problem is the area underneath $(1 - x)^3$ in the positive quadrant. This is shown in Figure 1. It is immediate to see that, under this constraint, the maximizer of out function is the point $(1, 0)$ - remember that we are trying to maximize the function $f(x) = x$.

Figure 1: Feasible set in problem 2



But now, look at what happens when we solve the problem using the Lagrangian method:

$$\mathcal{L} = x + \lambda \left[(1-x)^3 - y\right] + \mu_x x + \mu_y y$$

so that our KKT conditions are

$$\begin{cases} \mathcal{L}_x = 1 - 3\lambda(1-x^2) + \mu_x = 0 \\ \mathcal{L}_y = -\lambda + \mu_y = 0 \\ \lambda \geq 0, \mu_x \geq 0 \mu_y \geq 0 \\ \lambda \left[(1-x)^3 - y\right] = 0 \\ \mu_x x = 0, \mu_y y = 0 \end{cases}$$

What happens at $(1, 0)$? We have that $\mu_x = 0$ from CSC. But now look at the first FOC - since $\mu_x = 0 and x = 1$ we have $1 = 0$! The Lagrange method does not identify this maximizer.

Why? Because the CQ fails at $(1, 0)$. In fact, we have:

$$\begin{pmatrix} \nabla g_1(x,y) & \nabla g_2(x,y) & \nabla g_3(x,y) \end{pmatrix} = \begin{pmatrix} 3(1-x)^2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

At $(1, 0)$ the binding constraints are g_1 and g_3 and notice that $\nabla g_1(1, 0) = \begin{pmatrix} 0 & 1 \end{pmatrix}'$ so that the derivatives of the binding constraint are linearly dependent. In other words, CQ does not hold at $(1, 0)$.

Example 3 (Envelope Theorem)

Consider the following leisure-consumption problem. Here c represents consumption with unit price p, l represents leisure that can be substituted for labor at wage w, T represents the total amount of time allocated between labor and leisure and I represents outside income:

$$
\begin{aligned}
&\max c^{\alpha} \times l^{1-\alpha} \\
&\text{s.t } pc + wl \leq wT + I \\
&\quad l \leq T \\
&\quad c, l \geq 0
\end{aligned}
$$

The first thing to notice is that, at a maximum the non negativity constraints can not bind, else the objective function is zero, so we can focus on $c, l > 0$. Let us write the Lagrangian:

$$\mathcal{L} = c^{\alpha} \times l^{1-\alpha} + \lambda (wT + I - pc - wl) + \mu (T - l)$$

and KKT conditions are:

$$
\begin{cases}
\mathcal{L}_c = \alpha c^{\alpha} l^{1-\alpha} - \lambda p = 0 \\
\mathcal{L}_l = (1-\alpha) c^{\alpha} l^{-\alpha} - \lambda w - \mu = 0 \\
\lambda \geq 0, \mu \geq 0 \\
\lambda (wT + I - pc - wl) = 0 \\
\mu (T - l) = 0
\end{cases}
$$

It is also easy to se that the budget constraint must also bind at the optimum as one could always increase c for any given level of positive l and be better-off. Suppose that the only binding constraint is the budget constraint. We have:

$$
\begin{cases}
\lambda = \frac{\alpha l^{1-\alpha}}{pc^{1-\alpha}} = \frac{(1-\alpha)c^{\alpha}}{wl^{\alpha}} \\
wT + I = pc + wl
\end{cases}
$$

From these equations we get

$$
\begin{aligned}
c^* &= \frac{\alpha (wT + I)}{p} \\
l^* &= \frac{(1-\alpha)(wT + I)}{w} \lambda^* = \frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{p^{\alpha} w^{1-\alpha}}
\end{aligned}
$$

which is feasible as long as

$$l = \frac{(1-\alpha)(wT + I)}{w} < T \iff I < \frac{\alpha w T}{1-\alpha}$$

in which case this solution is not actually satisfying the leisure constraint. If the parameters fit this case, then the optimum actually occurs at $l = T$, and $c = I/p$, where both constraints are binding.

Now, suppose $I < \frac{\alpha w T}{1-\alpha}$, so the solutions are the ones we found above. Suppose we want to show that the indirect utility function satisfies the Envelope Theorem with respect to I.

We are gonna write $c^*(I)$ and $l^*(I)$ to underline that they depend on the parameter I, but we could do this with respect to any parameter. The Envelope theorem tells us that:

$$\frac{\partial f}{\partial I}(c^*(I), l^*(I)) = \frac{\partial \mathcal{L}}{\partial I}(c^*, l^*, \lambda^*, I)$$

Let's check this. First, we derive the value function $v(I) = f(c^*(I), l^*(I))$:

$$\begin{aligned} v(I) &= \left(\frac{\alpha\,(wT+I)}{p}\right)^{\alpha} \times \left(\frac{(1-\alpha)\,(wT+I)}{w}\right)^{1-\alpha} \\ &= \frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{p^{\alpha}w^{1-\alpha}} \times (wT+I) \end{aligned}$$

so that

$$\frac{\partial v}{\partial I} = \frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{p^{\alpha}w^{1-\alpha}}$$

We have found the first piece. Now consider:

$$\frac{\partial}{\partial I}\mathcal{L}(c^*, l^*, \lambda^*, I) = \lambda^* = \frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{p^{\alpha}w^{1-\alpha}}$$

so we have cheked that the Envelope theorem works with respect to I.

Example 4 (Envelope Theorem again)

Consider the problem

$$\max_{(x_1,x_2)\in\mathbb{R}^2_{++}} x_1^{\alpha}x_2^{1-\alpha}$$

s.t.

$$p_1x_1 + p_2x_2 \leq m$$

where $\alpha \in (0,1)$, $p_1, p_2 \in \mathbb{R}_{++}$, and $m \in \mathbb{R}_{++}$ are parameters. We ignore discussions on existence and applicability of KKT and go straight to the solution.

Write down the Lagrangian

$$L\,(x_1, x_2, \lambda) = x_1^{\alpha}x_2^{1-\alpha} + \lambda\,(m - p_1x_1 - p_2x_2)$$

and then the KKT condition

$$\begin{cases} x \in \mathbb{R}^2_{++} \\ \alpha x_1^{\alpha-1}x_2^{1-\alpha} - \lambda p_1 = 0 \\ (1-\alpha)\,x_1^{\alpha}x_2^{-\alpha} - \lambda p_2 = 0 \\ \lambda \geq 0,\ m - p_1x_1 - p_2x_2 \geq 0, \text{ and } \lambda\,(m - p_1x_1 - p_2x_2) = 0 \end{cases}$$

By the two FOCs, we have $\lambda > 0$, and so by CSC we have $m - p_1x_1 - p_2x_2 = 0$. Also, comparing the two FOCs gives us

$$\frac{\alpha x_1^{\alpha-1}x_2^{1-\alpha}}{(1-\alpha)\,x_1^{\alpha}x_2^{-\alpha}} = \frac{\lambda p_1}{\lambda p_2}$$

i.e.

$$\frac{p_1 x_1}{p_2 x_2} = \frac{\alpha}{1-\alpha}$$

and so we have

$$(x_1, x_2, \lambda) = \left(\frac{\alpha m}{p_1}, \frac{(1-\alpha)\, m}{p_2}, \frac{\alpha^{\alpha} (1-\alpha)^{1-\alpha}}{p_1^{\alpha} p_2^{1-\alpha}} \right)$$

as the unique solution to the K-T condition. So $(x_1, x_2) = (\alpha m / p_1, (1-\alpha)\, m / p_2)$ is the unique type 1 candidate in this problem.

Because CQ holds at all feasible point, there is no type 2 candidate at all. Because the problem has a solution by Weierstrass, we know that the unique type 1 candidate

$$(x_1, x_2) = \left(\frac{\alpha m}{p_1}, \frac{(1-\alpha)\, m}{p_2} \right)$$

must be the unique maximizer of the problem.

We can rewrite The solution to the Kuhn-Tucker conditions as

$$(x_1^*(p_1, p_2, m), x_2^*(p_1, p_2, m), \lambda^*(p_1, p_2, m)) = \left(\frac{\alpha m}{p_1}, \frac{(1-\alpha)\, m}{p_2}, \frac{\alpha^{\alpha} (1-\alpha)^{1-\alpha}}{p_1^{\alpha} p_2^{1-\alpha}} \right)$$

The value function of the maximization problem is

$$v(p_1, p_2, m) = (x_1^*)^{\alpha} (x_2^*)^{1-\alpha} = \frac{m \alpha^{\alpha} (1-\alpha)^{1-\alpha}}{p_1^{\alpha} p_2^{1-\alpha}}$$

Taking its first order derivative w.r.t. (p_1, p_2, m) we have:

$$\frac{\partial v}{\partial p_1} = -\frac{m \alpha^{1+\alpha} (1-\alpha)^{1-\alpha}}{p_1^{\alpha+1} p_2^{1-\alpha}} = -\lambda^* x_1^*$$

$$\frac{\partial v}{\partial p_2} = -\frac{m \alpha^{\alpha} (1-\alpha)^{2-\alpha}}{p_1^{\alpha} p_2^{2-\alpha}} = -\lambda^* x_2^*$$

$$\frac{\partial v}{\partial m} = \frac{\alpha^{\alpha} (1-\alpha)^{1-\alpha}}{p_1^{\alpha} p_2^{1-\alpha}} = \lambda^*$$

From the envelope theorem, we have:

$\frac{\partial \mathcal{L}}{\partial p_1}(x_1^*, x_2^*, \lambda^*, p_1, p_2, m) = -\lambda^* x_1^*,$

$\frac{\partial \mathcal{L}}{\partial p_2}(x_1^*, x_2^*, \lambda^*, p_1, p_2, m) = -\lambda^* x_2^*,$

$\frac{\partial \mathcal{L}}{\partial m}(x_1^*, x_2^*, \lambda^*, p_1, p_2, m) = \lambda^*.$