Equality constraints

1 Introduction

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{subject to } g_j(\mathbf{x}) = 0, \quad j = 1, 2, \ldots, m$$

The fundamental necessary conditions for an optimum are given by the Lagrange multiplier theorem. We consider three approaches:

- the perturbation (algebraic) approach
- the geometric approach
- the implicit function theorem approach.

2 Perturbation approach

2.1 Consumer problem

First, consider the consumer problem

$$\max_{\mathbf{x} \in X} u(\mathbf{x})$$
$$\text{subject to } \mathbf{p}^T\mathbf{x} = m$$

assuming that there are only two commodities ($X \subset \Re^2$). If $\mathbf{x}^*$ is optimal in the budget set

$$u(\mathbf{x}^* + \mathbf{dx}) \approx u(\mathbf{x}^*) + u_1 dx_1 + u_2 dx_2 \leq u(\mathbf{x}^*)$$

which implies that

$$u_1 dx_1 + u_2 dx_2 \leq 0$$

for all dx_1, dx_2 such that

$$p_1 dx_1 + p_2 dx_2 = 0$$

This requires that

$$\frac{u_1(\mathbf{x}^*)}{p_1} = \frac{u_2(\mathbf{x}^*)}{p_2} = \lambda$$
$$\textbf{or } \frac{u_1(\mathbf{x}^*)}{u_2(\mathbf{x}^*)} = \frac{p_1}{p_2}$$
$$\textbf{or } u_i(\mathbf{x}^*) = \lambda p_i, \qquad i = 1, 2$$
$$\textbf{or } \nabla u(\mathbf{x}) = \mathbf{p}$$

2.2 One constraint

Now consider the general optimization problem with a single equality constraint

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{subject to } g(\mathbf{x}) = 0$$

If $\mathbf{x}^*$ is optimal, we must have

$$f(\mathbf{x}^* + \mathbf{dx}) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)\mathbf{dx} \leq f(\mathbf{x}^*)$$

or

$$\nabla f(\mathbf{x}^*)\mathbf{dx} \leq 0$$

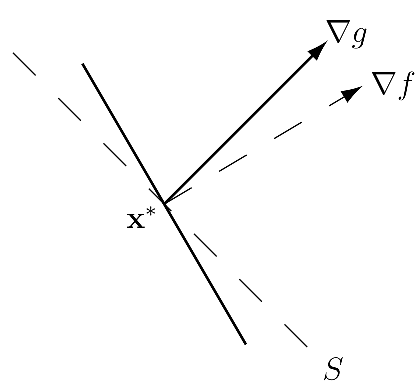
for all feasible perturbations $\mathbf{dx} \in X$ satisfying

$$\nabla g(\mathbf{x}^*)\mathbf{dx} = 0$$

which requires that

$$\nabla f(\mathbf{x}^*) = \lambda \nabla g(\mathbf{x}^*)$$

At a (local) optimum $\mathbf{x}^*$, the gradients of the objective function and the constraint must be aligned. Otherwise, it is possible to find a feasible perturbation that increases the value attained, as in the following diagram. Constraints limit the feasible perturbations that must be considered.



2.3 Multiple constraints

Generalizing this to multiple constraints is essentially an exercise in linear algebra. Suppose $\mathbf{x}^*$ is a local maximum in the general problem

$$\max_{\mathbf{x}} f(\mathbf{x})$$
$$\text{subject to } g_j(\mathbf{x}) = 0, \quad j = 1, 2, \ldots, m$$

If $\mathbf{x}^*$ is optimal, we must have

$$f(\mathbf{x}^* + \mathbf{dx}) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T \mathbf{dx} \leq f(\mathbf{x}^*)$$

or

$$\nabla f(\mathbf{x}^*)^T \mathbf{dx} \leq 0$$

for any feasible perturbation $\mathbf{dx} \in X$. The set of feasible (small) perturbation $\mathbf{dx}$ must satisfy

$$g_j(\mathbf{x}^* + \mathbf{dx}) = g_j(\mathbf{x}^*) + \nabla g_j(\mathbf{x}^*)^T \mathbf{dx} = 0$$

or

$$\nabla g_j(\mathbf{x}^*)^T \mathbf{dx} = 0, \quad j = 1, 2, \ldots, m$$

Let

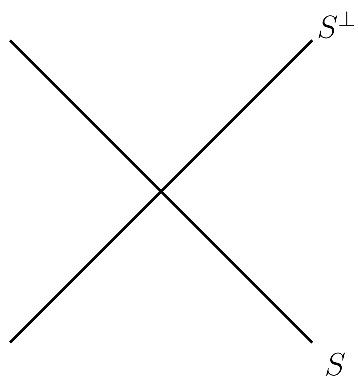
$$S = \{\mathbf{dx} : \nabla g_j(\mathbf{x}^*)^T \mathbf{dx} = 0, \quad j = 1, 2, \ldots, m\}$$

denote the set of feasible perturbations. Then S is a subspace, since S is the kernel of the linear function, the derivative of g. The orthogonal complement of S

$$S^\perp = \{\mathbf{a} \in \Re^n : \mathbf{a}^T \mathbf{dx} = 0, \mathbf{dx} \in S\}$$

is a subspace with

$$\nabla g_j(\mathbf{x}^*) \in S^\perp, j = 1, 2, \ldots, m$$



Since S is a subspace, optimality requires that

$$\nabla f(\mathbf{x}^*)^T \mathbf{dx} = 0 \text{ for every } \mathbf{dx} \in S$$

that is $\nabla f(\mathbf{x}^*)$ is orthogonal to the set of feasible perturbations S, or $\nabla f(\mathbf{x}^*) \in S^\perp$

Provided the gradients $\nabla g_1(\mathbf{x}^*), \nabla g_2(\mathbf{x}^*), \ldots, \nabla g_m(\mathbf{x}^*)$ are linearly independent, $Dg[\mathbf{x}^*]$ is of full rank m and its kernel S is a subspace of dimension

$n - m$. Therefore, $S^\perp$ is a subspace of dimension m, for which the m gradients $\nabla g_1(\mathbf{x}^*), \nabla g_2(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)$ provide a basis. Consequently, there exist unique coefficients $(\lambda_1, \lambda_2, \dots, \lambda_m)$ such that

$$\nabla f(\mathbf{x}^*) = \sum \lambda_j \nabla g_j(\mathbf{x}^*)$$

The coefficients λ_j are known as **Lagrange multipliers**. This result is an application of the **Fredholm alternative** (Exercises 3.48, 3.199, 3.237 and 3.238).

2.4 Lagrange multiplier theorem 5.2

Suppose that $\mathbf{x}^*$ is a local optimum of

$$\max_{\mathbf{x}} f(\mathbf{x})$$

$$\text{subject to } g_j(\mathbf{x}) = 0, \quad j = 1, 2, \dots, m$$

and $\nabla g_1(\mathbf{x}^*), \nabla g_2(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)$ are linearly independent. Then, there exist unique multipliers $\lambda_1, \lambda_2, \dots, \lambda_m$ such that

$$\nabla f(\mathbf{x}^*) = \sum \lambda_j \nabla g_j(\mathbf{x}^*)$$

2.5 Nonnegative variables (Corollary 5.2.1)

Suppose that $\mathbf{x}^*$ is a local optimum of

$$\max_{\mathbf{x} \geq \mathbf{0}} f(\mathbf{x})$$

$$\text{subject to } g_j(\mathbf{x}) = 0, \quad j = 1, 2, \dots, m$$

and $\nabla g_1(\mathbf{x}^*), \nabla g_2(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)$ are linearly independent. Then, there exist multipliers $\lambda_1, \lambda_2, \dots, \lambda_m$ such that

$$\nabla f(\mathbf{x}^*) \leq \sum \lambda_j \nabla g_j(\mathbf{x}^*) \quad \mathbf{x} \geq \mathbf{0} \quad \left(\nabla f(\mathbf{x}^*) - \sum \lambda_j \nabla g_j(\mathbf{x}^*)\right)^T \mathbf{x}^* = 0$$

Example (The consumer's problem)

$$\max_{\mathbf{x} \geq \mathbf{0}} u(\mathbf{x}) \text{ subject to } \mathbf{p}^T \mathbf{x} = m$$

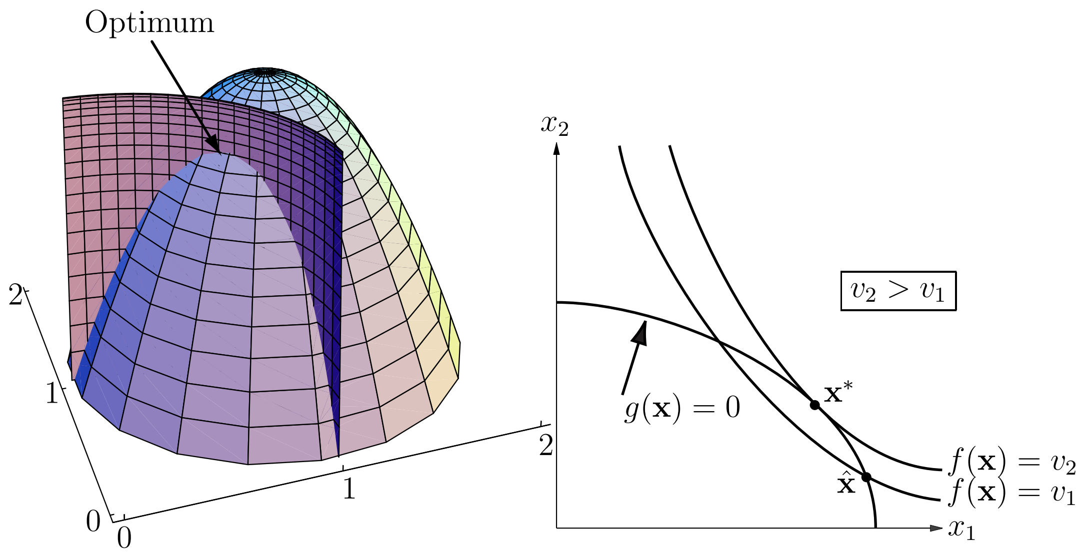
The first-order condition for (local) optimality is

$$\nabla u(\mathbf{x}^*) \leq \lambda \mathbf{p} \quad \mathbf{x} \geq \mathbf{0} \quad (\nabla u(\mathbf{x}^*) - \lambda \mathbf{p})^T \mathbf{x}^* = 0$$

That is, for every good i

$$u_i(\mathbf{x}^*) \leq \lambda p_i \text{ with } u_i(\mathbf{x}^*) = \lambda p_i \text{ if } x_i > 0$$

3 The geometric approach



Tangency between a contour of the objective function and the constraint is a necessary condition for optimality.

$$\text{Slope of } f = -\frac{f_1(\mathbf{x}^*)}{f_2(\mathbf{x}^*)} = -\frac{g_1(\mathbf{x}^*)}{g_2(\mathbf{x}^*)} = \text{Slope of } g$$

which equation implies

$$\frac{f_1(\mathbf{x}^*)}{g_1(\mathbf{x}^*)} = \frac{f_2(\mathbf{x}^*)}{g_2(\mathbf{x}^*)} = \lambda$$

which can be rewritten as two equations

$$f_1(x^*) = \lambda g_1(\mathbf{x}^*)$$
$$f_2(\mathbf{x}^*) = \lambda g_2(\mathbf{x}^*)$$

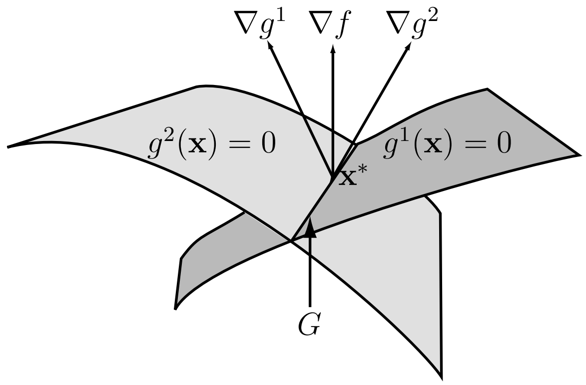
or more succinctly

$$\nabla f(\mathbf{x}^*) = \lambda \nabla g(\mathbf{x}^*)$$

This readily generalizes to

Many variables The constraint surface at $\mathbf{x}^*$ must be tangential to an indifference surface.

Multiple constraints Optimality requires that the gradient of the objective function be perpendicular (orthogonal) to the tangent plane of the constraint surface.

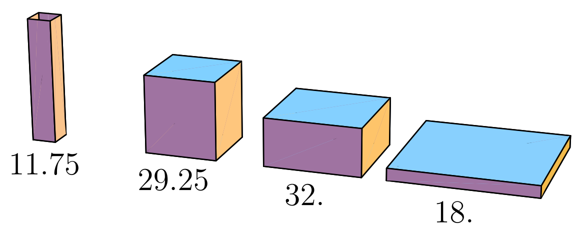


4 The implicit function theorem approach

4.1 A specific example

Design a rectangular vat (open at the top) to maximize the volume

$$\max w^2 h \text{ subject to } w^2 + 4wh = A$$



Solving the constraint for h

$$h = \frac{A - w^2}{4w}$$

and substituting into the objective function, we deduce

$$w^* = \sqrt{\frac{A}{3}} \text{ and } h^* = \frac{1}{2}\sqrt{\frac{A}{3}}$$

from which we can conclude that

$$h^* = \frac{1}{2} w^*$$

4.2 Consumer problem

To solve the consumer's problem

$$\max_{\mathbf{x}\in X} u(\mathbf{x})$$
$$\text{subject to } \mathbf{p}^T\mathbf{x} = m$$

we can solve the budget constraint for the first good

$$x_1 = \frac{m - \sum_{i=2}^n p_i x_i}{p_1}$$

to give an unconstrained maximization problem

$$\max u_{x_2,x_3,\dots,x_n}\Big(\frac{m - \sum_{i=2}^n p_i x_i}{p_1}, x_2, x_3, \dots, x_n\Big)$$

The first-order conditions for a maximum are

$$u_1\Big(-\frac{p_i}{p_1}\Big) + u_i = 0, \quad i = 2, 3, \dots, n$$

or

$$\frac{u_i}{p_i} = \frac{u_1}{p_1} = \lambda, \quad i = 2, 3, \dots, n$$

which is equivalent to

$$\nabla u(\mathbf{x}^*) = \lambda\mathbf{p}$$

4.3 General case

The implicit function theorem (Theorem 4.5) enables us to linearize the constraint

$$\mathbf{g}(\mathbf{x}) = \mathbf{0}$$

in a neighbourhood of $\mathbf{x}^*$, effectively solving for $x_1, x_2, \dots, x_m$ in terms of the remaining $n-m$ variables, converting the problem into one of unconstrained optimizaton.

This provides an entirely different derivation of the Lagrange Theorem 5.2, based on the implicit function theorem rather than the separating hyperplane theorem. Its insight and motivation are quite distinct.

5 The Lagrangean

5.1 Definition

The basic necessary condition

$$\nabla f(\mathbf{x}^*) = \sum \lambda_j \nabla g_j(\mathbf{x}^*)$$

can be expressed alternatively as

$$Df[\mathbf{x}^*] - \sum_{j=1}^{m} \lambda_j Dg_j[\mathbf{x}^*] = \mathbf{0}$$

which is the necessary condition for a stationary value of the function

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j g_j(\mathbf{x})$$

$L(\mathbf{x}, \boldsymbol{\lambda})$ is called the **Lagrangean** and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ the **Lagrange multipliers**.

5.2 Form of the Lagrangean

For equality constrained problems, it is immaterial whether we form the Lagrangean by adding or subtracting the constraints. The functions $L = f - \sum \lambda_j g_j$ and $L = f + \sum \lambda_j g_j$ will have the same stationary points, although the associated Lagrange multipliers will have opposite signs. However, with inequality constraints $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$, the chosen form $f - \sum \lambda_j g_j$ ensures that the Lagrange multipliers are nonnegative, which is the appropriate sign for their interpretation as shadow prices of the constraints (Section 6).

5.3 Necessary conditions (Corollary 5.2.2)

1. L stationary
2. L locally concave, but only with respect to the subspace of feasible perturbations

$$\mathbf{x}^T H_L(\mathbf{x}^*)\mathbf{x} \leq 0 \text{ for every } \ \mathbf{x} \in S = \{\mathbf{x} : Dg[\mathbf{x}^*](\mathbf{x}) = \mathbf{0}\}$$

$H_L(\mathbf{x}^*)$ is known as the bordered hessian (Simon and Blume 460-461, 467-468)

Example The point $(1, 1)$ is the only optimum of the problem

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

At the optimum, the Lagrange multiplier $\lambda^* = 1$ and the optimum $(1, 1)$ is in fact a saddle point of the Lagrangean

$$L(x_1, x_2, \lambda^*) = x_1 x_2 - (x_1 + x_2 - 2)$$

It maximizes the Lagrangean along the constraint $x_1 + x_2 = 2$.

5.4 Sufficient conditions (Corollary 5.2.3)

1. L stationary
2. L locally strictly concave, but only with respect to the subspace of feasible perturbations

$$\mathbf{x}^T H_L(\mathbf{x}^*)\mathbf{x} < 0 \text{ for every } \mathbf{x} \in S = \{\mathbf{x} : Dg[\mathbf{x}^*](\mathbf{x}) = \mathbf{0}\}$$

5.5 Global sufficient conditions

- L concave (Exercise 5.20)
  - in general knowledge of optimal λ is required before concavity can be verified
- f concave and $\mathbf{g}$ affine (Corollary 5.2.3)

5.6 Solving the first-order conditions

Compare Examples 5.22 and 5.23.

5.7 Constraint qualification

All of the above requires that the constraints be **regular** at $\mathbf{x}^*$, that is $D\mathbf{g}[\mathbf{x}^*]$ is of full rank or $\nabla g_1(\mathbf{x}^*), \nabla g_2(\mathbf{x}^*), \ldots, \nabla g_m(\mathbf{x}^*)$ are linearly independent. This is known as the **regularity** constraint qualification condition. More general conditions will be given later.

6 Shadow prices

In the consumer problem

$$du = \nabla u(\mathbf{x}^*)^T \mathbf{dx}$$

but

$$\mathbf{p}^T \mathbf{dx} = dM$$

and also

$$\nabla u(\mathbf{x}^*) = \mathbf{p}$$

Therefore

$$du = \lambda \mathbf{p}^T \mathbf{dx} = \lambda \, dM$$

In general

$$\max f(\mathbf{x}) \text{ subject to } g(\mathbf{x}) = c$$

If $c \to c + dc$ then $\mathbf{x}^* \to \mathbf{x}^* + \mathbf{dx}$ and

$$df = f(\mathbf{x}^* + \mathbf{dx}) - f(\mathbf{x}^*) \approx \nabla f(\mathbf{x}^*)^T \mathbf{dx}$$

But $\mathbf{dx}$ must satisfy

$$g(\mathbf{x} + \mathbf{dx}) \approx g(\mathbf{x}^*) + \nabla g(\mathbf{x}^*)^T \mathbf{dx} = c + dc$$

$$\nabla g(\mathbf{x}^*)^T \mathbf{dx} = dc$$

$\mathbf{x}^*$ satisfies the first-order condition

$$\nabla f(\mathbf{x}^*) = \lambda \nabla g(\mathbf{x}^*)$$

Substituting

$$df = \nabla f(\mathbf{x}^*)^T \mathbf{dx} = \lambda \nabla g(\mathbf{x}^*)^T \mathbf{dx} = \lambda dc$$

or

$$\frac{\partial f}{\partial c} = \lambda$$

The Lagrange multiplier λ measures the **shadow price** of the constraint.

Example In the vat design problem, $\lambda = 1/2$. Any additional sheet metal is worth approximately one half a cubic meter of volume. Suppose that volume is worth p per cubic meter. Then $p\lambda$ is the shadow price of sheet metal in dollars. If additional sheet metal can be bought at price q, then purchase is worthwhile provided its shadow price $p\lambda$ exceeds its market price q.

7 Homework

1. Characterize the optimal solution of the general two variable constrained maximization problem

$$\max_{x_1, x_2} f(x_1, x_2)$$
$$\text{subject to } g(x_1, x_2) = 0$$

   using the implicit function theorem to solve the constraint.

2. Solve the general Cobb-Douglas utility maximization problem

$$\max_{\mathbf{x}} u(\mathbf{x}) = x_1^{\alpha_1} \alpha_2^{\alpha_2} \ldots\, x_n^{\alpha_n}$$
$$\text{subject to } p_1 x_1 + p_2 x_2 + \ldots p_n x_n = m$$

   [Hint: Follow the technique in Example 5.23.]

3. Assume a consumer has a quasi-linear preferences

$$u(\mathbf{x}) = x_1 + a \log x_2$$

   Solve the consumer's problem ensuring that consumption is nonnegative. For simplicity, assume that $p_1 = 1$.

4. Suppose that $(\mathbf{x}^*, \boldsymbol{\lambda})$ is a stationary point of the Lagrangean

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum \lambda_j g_j(\mathbf{x})$$

   and $L(\mathbf{x}, \boldsymbol{\lambda})$ is concave in $\mathbf{x}$. Then $\mathbf{x}^*$ is a global solution of the problem

$$\max_{\mathbf{x} \in X} f(\mathbf{x}) \text{ subject to } \mathbf{g}(\mathbf{x}) = \mathbf{0}$$

Solutions: Equality constraints

1 Assume that $\mathbf{x}^* = (x_1^*, x_2^*)$ solves

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

subject to

$$g(x_1, x_2) = 0$$

By the implicit function theorem, there exists a function $h : \Re \to \Re$ such that

$$x_1 = h(x_2) \tag{1}$$

and

$$g(h(x_2), x_2) = 0$$

for x_2 in a neighborhood of x_2^*. Furthermore

$$Dh[x_2^*] = -\frac{D_{x_1} g[\mathbf{x}^*]}{D_{x_2} g[\mathbf{x}^*]} \tag{2}$$

Using (1), we can convert the original problem into the unconstrained maximization of a function of a single variable

$$\max_{x_2} f(h(x_2), x_2)$$

If x_2^* maximizes this function, it must satisfy the first-order condition (applying the Chain Rule)

$$D_{x_1} f[x*] \circ Dh[x_2^*] + D_{x_2} f[\mathbf{x}^*] = 0$$

Substituting (2) yields

$$D_{x_1} f[x*] \left(-\frac{D_{x_1} g[\mathbf{x}^*]}{D_{x_2} g[\mathbf{x}^*]} \right) + D_{x_2} f[\mathbf{x}^*] = 0$$

or

$$\frac{D_{x_1} f[x*]}{D_{x_2} f[\mathbf{x}^*]} = \frac{D_{x_1} g[\mathbf{x}^*]}{D_{x_2} g[\mathbf{x}^*]}$$

2 The Lagrangean is

$$L(\mathbf{x}, \lambda) = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} - \lambda(p_1 x_1 + p_2 x_2 + \dots + p_n x_n)$$

The first-order conditions for a maximum are

$$D_{x_i} L = \alpha_i x_1^{\alpha_1} x_2^{\alpha_2} \dots x_i^{\alpha_i - 1} \dots x_n^{\alpha_n} - \lambda p_i = \frac{\alpha_i u(x)}{x_i} - \lambda p_i = 0$$

or

$$\frac{\alpha_i u(x)}{\lambda} = p_i x_i \qquad i = 1, 2, \ldots, n \tag{3}$$

Summing over all goods and using the budget constraint

$$\sum_{i=1}^n \frac{\alpha_i u(x)}{\lambda} = \frac{u(x)}{\lambda} \sum_{i=1}^n \alpha_i = \sum_{i=1}^n p_i x_i = m$$

Letting $\sum_{i=1}^n \alpha_i = \alpha$, this implies

$$\frac{u(\mathbf{x})}{\lambda} = \frac{m}{\alpha}$$

Substituting in (3)

$$p_i x_i = \frac{\alpha_i}{\alpha} m$$

or

$$x_i^* = \frac{\alpha_i}{\alpha} \frac{m}{p_i} \qquad i = 1, 2, \ldots, n$$

3 The consumer's problem is

$$\max_{\mathbf{x} \geq 0} u(\mathbf{x}) = x_1 + a \log x_2$$
$$\text{subject to } g(\mathbf{x}) = x_1 + p_2 x_2 - m = 0$$

The first-order conditions for a (local) optimum are

$$D_{x_1} u(\mathbf{x}^*) = 1 \leq \lambda = D_{x_1} g(x^*) \qquad x_1 \geq 0 \qquad x_1(1 - \lambda) = 0 \tag{4}$$

$$D_{x_2} u(\mathbf{x}^*) = \frac{a}{x_2} \leq \lambda p_2 = D_{x_2} g(x^*) \qquad x_2 \geq 0 \qquad x_2 \left(\frac{a}{x_2} - \lambda p_2 \right) = 0 \tag{5}$$

We can distinguish two cases:

Case 1 $x_1 = 0$ in which case the budget constraint implies that $x_2 = m/p_2$.

Case 2 $x_1 > 0$ In this case, (4) implies that $\lambda = 1$. Consequently, the first inequality of (5) implies that $x_2 > 0$ and therefore the last equation implies $x_2 = a/p_2$ with $x_1 = m - a$.

We deduce that the consumer first spends portion a of her income on good 2 and the remainder on good 1.

4 Corollary 5.1.2 implies that $\mathbf{x}^*$ is a global maximum of $L(\mathbf{x}, \boldsymbol{\lambda})$, that is

$$L(\mathbf{x}^*, \boldsymbol{\lambda}) \geq L(\mathbf{x}, \boldsymbol{\lambda}) \text{ for every } \mathbf{x} \in X$$

which implies

$$f(\mathbf{x}^*) - \sum \lambda_j g_j(\mathbf{x}^*) \geq f(\mathbf{x}) - \sum \lambda_j g_j(\mathbf{x}) \text{ for every } \mathbf{x} \in X$$

Since $\mathbf{g}(\mathbf{x}^*) = \mathbf{0}$ this implies

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) - \sum \lambda_j g_j(\mathbf{x}) \text{ for every } \mathbf{x} \in X$$

A fortiori

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{ for every } \mathbf{x} \in G = \{\, \mathbf{x} \in X : \mathbf{g}(\mathbf{x}) = \mathbf{0} \,\}$$