LN - 3.0 Constrained optimization: the Lagrangean method

3.1. Existence of a solution to the UMP
3.2. Utility maximization with only the wealth constraint
   3.2.1 First order necessary conditions
   3.2.2 Second order necessary conditions
3.3. Maximization with only the nonnegativity constraints on the variables
3.4. Constrained optimization with wealth and nonnegativity constraints
3.5. The expenditure minimization problem (EMP)
3.6 The role of concavity and quasiconcavity in optimization problems

Appendix A.3 The analytics of the utility maximization problem
Appendix A.3.1 Unconstrained optimization
Appendix A.3.2 Constrained optimization: Lagrange’s theorem and function
Appendix A.3.3 Nonnegativity constrained optimization: Kuhn-Tucker conditions
Appendix A.3.4 Inequality constrained optimization: the saddle point solution

The principal aim of this Lecture Note lies in the correct formulation of constrained optimization problems and in the economic interpretation of some auxiliary variables that are integral part of the solution. In line with the development of classical demand theory, we focus on the utility maximization problem and leave to the end of the Lecture Note a presentation of the expenditure minimization problem.¹

The analytical tool for the determination of the solution of constrained optimization problems consists in the definition of the Lagrangean function associated to the problem and thus in solving the original constrained optimization as an unconstrained maximization or minimization problem, as the case may be. The Lagrangean function introduces in the analytical formulation of the problem some auxiliary variables – the Lagrangean multipliers. These variables are shadow prices, which implicitly measure the role played by the constraints in determining the optimum value of the objective function and, consequently, the implications of softening or tightening them. The Kuhn-Tucker theorem extends Lagrange’s

¹ An equally typical optimization problem, the cost of production minimization problem, will be examined in the Lecture Note on the theory of the firm.
solution of the classical equality constrained optimization problem to problems of inequality constrained optimization.

The utility maximization problem, henceforth UMP, represents the thread of the presentation through sections 3.1-3.4. We start in Section 3.1 with the analytical formulation of the UMP subject to the budget constraint and to the nonnegativity constraint on the consumption of the various commodities and establish the conditions for the existence of a solution. The UMP subject only to the wealth constraint, disregarding therefore the nonnegativity constraint, is examined in Section 3.2, while section 3.3 introduces, with reference to a function of a single variable, the problem posed by the nonnegativity constraints and explains the Kuhn-Tucker compact formulation of the conditions for the determination of the critical values of the objective function. The complete solution of the UMP, stated in Section 3.1, completes the study in Section 3.4. The Lagrangean formulation of the expenditure minimization problem, henceforth EMP, and the interpretation of the multipliers conclude the Lecture in Section 3.5.

The mathematical Appendix offers a generalization of the theory of constrained optimization to the case of several constraints. The analytical aspects of the Lagrange’s and the Kuhn-Tucker theorems are examined in Section 3.A.2 and its various subsections. It is preceded in Section 3.A.1 by a review of the distinction between second order necessary and sufficient conditions for a maximum of an unconstrained one variable function.

3.1. Existence of a solution to the UMP

Let \( u(x) \) be a continuous, twice differentiable utility function representing convex, monotone and continuous preferences defined on the nonnegative orthant of the commodity space \( \mathbb{R}^L_+ \). Let \( p = (p_1, p_2, \ldots, p_L) \) be the strictly positive vector of market prices of the \( L \) commodities, at which the consumer can buy any quantity of any good,\(^2\) and \( w \) the positive wealth of the consumer.\(^3\) The utility maximization problem can be compactly formulated in the following analytical terms

\[
\max_{x} u(x) \quad \text{subject to } p \cdot x - w \leq 0 \text{ and } x \geq 0
\]

(3.1)

where, in view of the following study of the solution of the problem, we have separately indicated the constraints represented by the wealth endowment of the consumer – for short,

\( \phantom{\text{3.1. Existence of a solution to the UMP}} \)

\(^2\) This means that the consumer operates in competitive commodity markets.

\(^3\) Following MWG’s approach, the resource constraint is specified in terms of wealth rather than in the more usual terms of income. The term wealth makes it possible to accommodate in a unique notation also problems of intertemporal choice in which the consumer’s resources are determined by his life cycle wealth.
the wealth constraint – and the nonnegativity requirement on the variables. These two constraints merge in the definition of the consumer budget set

\[ B(p, w) = \{ x \in \mathbb{R}_+^d \mid p \cdot x - w \leq 0 \} \]

The outward boundary of the budget set is the budget line \( p \cdot x - w = 0 \).

Whereas the assumptions on the utility function are more stringent than necessary for the existence of a solution, the assumption concerning prices and wealth are necessary. The assumption that wealth is positive is required in order to have a meaningful solution, for otherwise the only solution would the uninteresting null vector \( x = 0 \). The assumption that all prices are strictly positive is crucial for the boundedness of the budget set.

With prices and wealth all positive, the budget set is the closed and bounded shaded area in Fig. 3.1 with intercepts \( \left( \frac{w}{p_1}, \frac{w}{p_2} \right) \). Consider now, for contradiction, the sequence \( p_i^n \to 0 \) with \( p_2 \) and \( w \) both positive and fixed. The budget line rotates outward around point \( A \), as indicated by the arrow in the diagram; the intercept on the \( x_1 \)-axis moves further and further out. As long as \( p_i^n \) is positive, the intercept remains finite and the budget set is bounded. At the limit, however, when \( p_i^n \) is equal to zero, the intercept on the \( x_1 \)-axis goes to infinity and the budget set becomes the unbounded set \( \{ x_1 \in [0, \infty), x_2 \in [0, \frac{w}{p_2}] \} \).

Differentiability, in particular, is not required.
Fig. 3.1 – Unboundedness of the budget set when \( p_t \) tends to zero

We conclude that the prices of all commodities must be positive in order that the budget set be bounded and, therefore, compact. We can therefore assert that the Weierstrass Extreme Value Theorem is applicable to our utility maximization problem since, by assumption, \( u(x) \) is continuous and the budget set \( B(p,w) \) is non empty, convex and compact.\(^5\) This establishes the existence of a solution.

3.2. Utility maximization with only the wealth constraint

We begin with the simplest constrained optimization problem, that of maximizing the utility function \( u(x) \) subject only to the wealth constraint. We suppose, in other words, that the nonnegativity constraint on the variables may be, for the moment, disregarded on the tacit assumption that the UMP has an internal solution, which excludes the possibility that the quantity of one of the commodities is equal to zero. Taking furthermore into account the assumption of monotone preferences, we can write the wealth constraint with the equality sign, since the maximizing solution must necessarily occur on the boundary of the budget set, that is on the budget line.

With these assumptions, the general utility maximization problem (3.1) reduces to the following simpler problem

\[
\begin{align*}
\max_{x} & \quad u(x) \\
\text{subject to} & \quad p \cdot x - w = 0
\end{align*}
\] (3.4)

The analytical techniques for determining the solution of an optimization problem involving a twice continuously differentiable function, say \( f(x) \), defined in the open domain \( \text{int} \ D \) a subset of \( \mathbb{R}^L_+ \), whether subject to constraints or not, goes through a first step consisting in verifying the conditions for the existence of a solution, and a subsequent step consisting in the determination of the necessary and sufficient conditions for an optimum, a maximum or a minimum of \( f(x) \).

We have already taken care of the first step, namely of the conditions for the application of Weierstrass Extreme Value Theorem to the UMP, and concentrate, therefore, in this and in

\(^5\) Convexity of the budget set follows from the assumption that consumers operate in competitive markets where non linear pricing policies are excluded
the following sections on the second step that involves the determination of the first and second order derivatives of \( f(x) \). The first order conditions identify the critical points of the function as the vector that solves the system of equations obtained by setting the first partial derivative of the function equal to zero. The critical points may be maxima or minima or saddle points. The second order condition distinguishes among these according to the properties of the Hessian matrix of the second order derivatives, possibly on a proper subspace of the domain.

### 3.2.1 First order necessary conditions

The standard approach to the solution of this optimization problem is to use the Lagrangean method. The same result can actually be obtained by incorporating the wealth constraint in the objective function, we call this for short the “substitution approach” as opposed to the Lagrangean approach. We pursue first this alternative road, which has the advantage of clarifying the connection between the second order conditions for a maximum and the property of the utility function, and return subsequently to the formulation of the Lagrangean function.

Solving the constraint, for instance, for \( x_L \), we have

\[
x_L = x_L (x_1, \ldots, x_{L-1}; w) = \frac{w}{p_L} \sum_{i=1}^{L-1} p_i x_i
\]

Substituting for \( x_L \) from (3.5) in \( u(x) \) we obtain a new function of only \( L-1 \):

\[
h(x) = h(x_1, \ldots, x_{L-1}, x_L (x_1, \ldots, x_{L-1}; w)) = h(x_1, \ldots, x_{L-1})
\]

Taking derivatives with respect to the \( L-1 \) commodities, we obtain, for the generic commodity \( l \), the following first order condition for a critical value of \( v(x) \)

\[
\frac{\partial v(x_1, \ldots, x_{L-1})}{\partial x_l} = \frac{\partial u(x)}{\partial x_l} + \frac{\partial u(x)}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial u(x)}{\partial x_l} + \frac{\partial u(x)}{\partial x_L} \left( \frac{p_l}{p_L} \right) = 0
\]

where \( x^* \) is the solution vector. We can rewrite (3.6) as

\[
MRS_{u,l} = \frac{\partial u(x^*)}{\partial x_l}/ \frac{\partial u(x^*)}{\partial x_L} = \frac{p_l}{p_L} \quad \text{for all} \quad l=1,\ldots,L-1
\]

This is a set of \( L-1 \) equations which impose the condition that the marginal rates of substitution with respect to commodity \( L \) be equal to the corresponding price ratios, i.e. a condition of tangency between the indifference curves of the subsets of the two commodities \((l,L)\) and the corresponding price ratios, as depicted in Fig. 3.2 with \( l=1 \) and \( L=2 \). Given
the assumptions made at the beginning of Section 3.1, strengthened to strict convexity of preferences, the system of equations (3.7) has a unique solution; substituting in (3.5) also $x_L$ can be determined.

Consider now Lagrange’s method of analyzing the problem of utility maximization under constraint, which consists in transforming the constrained optimization problem (3.4) into an unconstrained one by adding a new variable the Lagrangean multiplier $\lambda$. In setting up the Lagrange’s function care must be taken that, in the solution, the multiplier be of the right sign: it should indicate the direction of change of the value function of the problem resulting from an increase of wealth, i.e. from a relaxation of the constraint. To this end, we follow the convention to set up the constraint in the “less than or equal” form and the multiplier with a minus sign; alternatively, the constraint can be equivalently written in the “greater than or equal” form and the multiplier with a plus sign. The Lagrangean is accordingly

$$L(x, \lambda) = u(x) - \lambda (p \cdot x - w)$$

The critical points of the Lagrangean are determined by the condition that the partial derivatives with respect to the variables $x$ and $\lambda$ be equal to zero. We have

$$\frac{\partial L(x, \lambda)}{\partial x} = \nabla u(x^*) - \lambda^* p = 0$$
$$\frac{\partial L(x, \lambda)}{\partial \lambda} = p \cdot x^* - w = 0$$

where $(x^*, \lambda^*)$ is the solution of this set of equations.

Lagrange’s theorem establishes that, if $x^*$ maximizes $u(x)$ subject to the wealth constraint, than there exists a unique number $\lambda^*$ that satisfies the set of equations in the first line of (3.9). Note that $\lambda^*$ is certainly positive because, even in the case, to be later examined, of a boundary solution and given that we have assumed that all prices and wealth are strictly positive, at least one of the elements of the vector $x^*$ must be positive, so that the corresponding condition in the first line of (3.9) must be satisfied with the equal sign.

The economic meaning of the multiplier $\lambda$ follows directly from the fact that wealth represents the binding constraint on the utility level that can be achieved. A softening of the

---

6 Lagrange’s theory of constrained optimization is contained in his Lecons sur le calcul des functions (1806).
7 As we will argue in the Appendix A.3.2, it would not be correct to say that the solutions values $(x^*, \lambda^*)$ are the maximizers of the Lagrangean function (3.8). For a proof of Lagrange’s Theorem see Appendix 3.A. See also MWG, pp. 956-7 and JR, pp. 577-601. A very simple proof for the two-commodity case is offered by Simon-Blume, pp. 413-415.
constraint, that is an increase in wealth, makes it possible to reach a higher indifference curve, as in Fig. 3.2. The positive value of $\lambda^*$ measures, therefore, the marginal increase in utility that the consumer can attain thanks to a differential increase in wealth, namely the marginal utility of wealth. A word of caution about the numerical value of $\lambda^*$ is in order. Since the utility function representing preferences is defined up to an increasing monotonic transformation, $\lambda^*$ measures the marginal utility of wealth with reference to the specific utility function chosen in the class of function representing given preferences.

![Fig. 3.2 – Effect of a wealth increase](image)

**3.2.2 Second order necessary conditions**

Let us turn to the second order necessary condition for a maximum. Following the approach already used, we first consider the “substitution approach” to utility maximization. Differently from what we have done in the preceding Section, we focus, however, for ease of derivation on the two commodity case; the extension to the case of more than two commodities is straightforward.

---

9 As shown in appendix A.3.2, this is in effect the minimum value assigned to the marginal utility of wealth.
Substituting for commodity 2 in terms of commodity 1, as indicated in (3.5) above, in the utility function \( u(x) \), we can reduce the problem of maximization under constraint to the problem of unconstrained maximization of the following function of the single variable \( x_1 \)

\[
(3.10) \quad h(x_1) = u\left(x_1, \frac{w}{p_2} - \frac{p_1}{p_2}x_1 \right)
\]

Taking first and second order derivatives of \( h(x_1) \) we have

\[
(3.11) \quad h'(x_1) = u_1(x_1) + u_2(x_1)\left(-\frac{p_1}{p_2}\right)
\]

\[
(3.12) \quad h''(x_1) = u_{11}(x_1) + u_{12}(x_1)\left(-\frac{p_1}{p_2}\right) + \left[u_{21}(x_1) + u_{22}(x_1)\left(-\frac{p_1}{p_2}\right)\right] - \frac{p_1}{p_2} =
\]

\[
= \left(\frac{1}{p_2}\right)^2 \left[p_1^2 u_{11} - 2p_1 p_2 u_{12} + p_2^2 u_{22}\right] = \left(\frac{1}{p_2}\right)^2 \left[-\det \tilde{H}^B(x^*)\right]
\]

where first and second order partial derivatives of \( u(x) \) refer in proper order to the derivatives with respect to the first and the second argument of the function \( h(x_1) \) and \( \tilde{H}^B(x^*) \) is the 3x3 matrix in which the Hessian matrix of the utility function \( u(x) \) is bordered with the price vector

\[
(3.13) \quad \tilde{H}^B(x^*) = \begin{bmatrix}
    u_{11} & u_{12} & p_1 \\
    u_{21} & u_{22} & p_2 \\
    p_1 & p_2 & 0
\end{bmatrix}
\]

As will be shown in the Appendix, Section A.3.2 this is the Bordered hessian of the Lagrange’s function.

The first and second order necessary conditions for a maximum are then, respectively, \( h'(x_1) = 0 \), which implies the equality between the marginal rate of substitution and the price ratio

\[
(3.14) \quad \frac{u_1(x)}{u_2(x)} = \frac{p_1}{p_2}
\]

---

10 The approach followed here parallels the technique used in Lecture Note 2 to exemplify the notion that a negative semidefinite matrix of a function of several variables is the analogue of a nonpositive second order derivative for a function of a single variable.
and $h^*(x_i) \leq 0$, which implies that $h^*(x_i)$ is negative (nonpositive) if and only if $\tilde{H}^B(x^*)$ is negative semidefinite, that is if $\det \tilde{H}^B(x^*)$ is positive (nonnegative).\footnote{See Lecture Note 2, Section 2.3.B.1.} This last statement means that $H(x)$ is negative semidefinite in the linear space defined by the wealth constraint.

This conclusion establishes a direct connection between the second order necessary conditions for a maximum and the properties of the utility function. In order to better pursue this connection, let us first note that the matrix $\tilde{H}^B(x^*)$ is similar to, but is not the Bordered Hessian $H^B(x^*)$ of the utility function $u(x)$, since the borders of the matrix $\tilde{H}^B(x^*)$ are not the partial derivatives of $u(x)$. In fact $H^B(x^*)$ is

$$H^B(x) = \begin{bmatrix} u_{11} & u_{12} & u_1 \\ u_{21} & u_{22} & u_2 \\ u_1 & u_2 & 0 \end{bmatrix}$$

But note - and this is the relevant point here - that given the direct proportionality between the price vector $p$ and the vector of partial derivative $\Delta u(x^*)$, the determinant of $\tilde{H}^B(x^*)$ is just proportional to the determinant of $H^B(x^*)$.

As we have seen in Lecture Note 2, the utility function $u(x)$ is quasiconcave if the Hessian matrix of second order derivatives of $u(x)$ is negative semidefinite in the linear subspace $Z = \{ z \in \mathbb{R}_+^L \mid \nabla u(x) \cdot z = 0 \}$. To grasp the implications of this condition, we make as usual reference to the two variable case and reproduce below Fig. 2.7 of Lecture Note 2, renumbered as Fig. 3.3.
The line $AB$ is now the wealth constraint $p \cdot x^* - w = 0$ tangent to the highest attainable level curve of $u(x)$ at the point $x^*$. Consider an arbitrary small change $\Delta x$ about $x^*$, with the vector $\Delta x \in \mathbb{R}^2$ and $x^* + \Delta x \in \mathbb{R}^2_{++}$, satisfying the wealth constraint, so that we have $p \cdot \Delta x = 0$. Substituting the first order conditions for a maximum $p = (\lambda^*)^{-1} \Delta u(x^*)$ and, eliminating $\lambda^* > 0$, we have $\nabla u(x^*) \cdot \Delta x = 0$. Noting that $\Delta x$ stands for the vector $z$ in the Definition 2.10 of a quasiconcave function given in Lecture Note 2 the connection between the definition of quasiconcavity of the utility function and the second order necessary condition for a maximum is established.\(^{12}\)

The second order sufficient condition for a maximum is $h''(x_i) < 0$ for $x_i \neq 0$. The difference between second order necessary and sufficient conditions is illustrated in the appendix A.3.1 while the analytics of the determination of the second order conditions for a maximum following the Lagrangean approach is postponed to Appendix A.3.2.

Appendix A.3.1 provides an analytical explanation of the role of the first and second order necessary conditions in the determination of the maximizers of the utility function. We look here for an intuitive explanation using economic considerations that can be easily drawn from the diagram of Fig. 3.2. With $u(x)$ representing monotone and convex preferences, it is obvious that an internal solution, as assumed, must be represented by the tangency point between the budget set and the highest indifference curve attained on the budget line.\(^{13}\) Points

---

\(^{12}\) This connection is further explored in Section 3.6.

\(^{13}\) We disregard here the possibility that the tangency point may occur on the coordinate axes.
below the budget line are not maximizers, points above are unfeasible. The motivation for the condition that the determinant of the bordered matrix \( \tilde{H}^B(\tilde{x}^*) \), involving the Hessian of the utility function, be negative semidefinite is, as explained in Lecture Note 2, a generalization to the case of a vector variable of the condition of a non positive second order derivative in the case of a scalar variable. Fig. 3.2 provides again the required intuition. A departure from the tangency point, in the typical case of a strictly convex indifference curve, causes a movement to a lower indifference curve and, therefore, a reduction in the utility level that could otherwise be attained.

3.3. Maximization with only the nonnegativity constraint on the variables

Our next step is to study the problem of determining the conditions for the maximization of a continuous twice differentiable function \( f(x) \) assuming that the only constraints on the problem are represented by the nonnegativity condition on the vector of the choice variables \( x \). The problem is best examined if we consider that \( f(x) \) is a function of a single variable and thus consider \( x \) a scalar. This has the advantage of explaining most clearly how we should deal with a nonnegativity constraint, which may be binding or not, and of easily deriving the consequent compact formulation of the first order conditions known as the Kuhn-Tucker conditions. Although not formally specified by the Kuhn-Tucker conditions, the economic meaning of the Lagrange’s multipliers of the nonnegativity constraints can be easily identified. Again, the generalization to several variables does not present analytical difficulties.

The maximization problem is now the following

\[
\begin{align*}
\max_{x} & \quad f(x) \\
\text{subject to} & \quad x \geq 0
\end{align*}
\]

Fig. 3.4 shows the three possible situations that may arise. In Panel (a) the constraint is not binding: the function \( f(x) \) attains its maximum at \( x^* > 0 \) inside the feasible set; the interior solution is then determined by the standard condition for an unconstrained problem \( f'(x^*) = 0 \). In Panels (b) and (c) the optimal solution is on the boundary of the constraint set, at \( x^* = 0 \). In Panel (b), the standard first order condition \( f'(x^*) = 0 \) obtains, while in Panel (c) at \( x^* = 0 \) is associated the condition \( f'(x^*) < 0 \).
Fig. 3.4 – Three possible maximizations with a nonnegativity constraint

An important fact emerges: in all three instances two conditions characterize now the optimal solution: the first regarding the first order derivative, the second the value of the maximer. Summing up in the order of three panels of Fig. 3.4, we have

\[
\begin{cases}
\text{Panel (a)} & f'(x^*) = 0, x^* > 0 \\
\text{Panel (b)} & f'(x^*) = 0, x^* = 0 \\
\text{Panel (c)} & f'(x^*) < 0, x^* = 0 \\
\end{cases}
\]

(3.15)

It is evident that these three cases can be synthesized into the following optimum conditions, the Kuhn-Tucker conditions

\[
\begin{align*}
& f'(x^*) \leq 0 \\
& x^* f'(x^*) = 0 \\
& x^* \geq 0
\end{align*}
\]

(3.16)

or, disregarding the implicit condition that the optimal solution is nonnegative, in more compact form

\[
\begin{align*}
& f'(x^*) \leq 0 \quad \text{with equality if } x^* > 0 \\
& x^* f'(x^*) = 0
\end{align*}
\]

(3.17)

We will for short, albeit somewhat improperly, refer in particular to the condition \( x^* f'(x^*) = 0 \) in the second line of (3.17), which synthesizes the crucial aspect concerning the presence of the nonnegativity constraints, as the Kuhn-Tucker condition.\(^{14}\) This is called the complementary slackness condition.

\(^{14}\) Economists traditionally refer to the conditions of inequality constrained optimization problems – and, in particular, to problems characterized by nonnegativity constraints on the choice variables – as the Kuhn-Tucker conditions, so
The Lagrange’s function for the maximization problem (3.14) can now formulated following the convention stated in the preceding Section on the utility maximization problem subject only to the wealth constraint. Thus, also the nonnegativity the constraint must be set up in the “less than or equal” form and the multiplier preceded by a minus sign. Using the Greek letter $\mu$ for the multiplier of the nonnegativity constraint, the Lagrange’s function is

$$L(x, \mu) = f(x) - \mu(-x)$$  \hspace{1cm} (3.18)

The first order conditions for a critical value of $L(x, \mu)$ with respect to the two variables $x$ and $\mu$ are

$$\frac{\partial L(x, \mu)}{\partial x} = f'(x^*) + \mu^* = 0$$  \hspace{1cm} (3.19)

$$\frac{\partial L(x, \mu)}{\partial \mu} = x \geq 0 \text{ with the added slack condition } \mu^* = 0 \text{ if } x^* > 0, \mu^* \geq 0 \text{ if } x^* = 0$$

The added slack conditions can be formulated in the compact form

$$\mu^* x^* = 0$$  \hspace{1cm} (3.20)

which covers the three possible case of relation (3.15).

In the Kuhn-Tucker conditions (3.16) the multiplier $\mu^*$ does not formally appear, but is clearly implied by it. The Lagrangean formulation of the optimization problem subject to the nonnegativity constraint clarifies the meaning of the expression complementary slack condition. It states that if the constraint is slack, i.e. non binding and thus implying $x^* > 0$, the multiplier must $\mu^*$ be zero; if, on the contrary, the Lagrangean multiplier is strictly positive, the constraint must be binding. This is exactly what the first order conditions (3.19) for a critical value of $L(x, \mu)$ say. Suppose that $x^* > 0$. Then from the second line of (3.19) we know that $\mu^* = 0$ and from the first line we have, therefore, $f(x^*) = 0$. If, instead, the constraint is strictly binding, then $\mu^* > 0$ and in this case the first line of (3.19) is satisfied with $f(x^*) < 0$. The meaning of the positive value of the multiplier in this latter case is easily explained looking at Fig. 3.5. If the constraint were relaxed and be, for example, $x \geq -1$, the function $f(x)$ could achieve a higher value. Thus, also in this case, the value of the

---

named after their paper published in 1951. A more correct reference should actually be to Karush-Kuhn-Tucker (KKT) conditions, in line with later scholars’ discovery that the first order necessary conditions for the solution of this type of optimization problems had been already stated by W. Karush in 1939.
multiplier measures the benefit in terms of the optimal value of the objective function deriving from a relaxation of the constraint.

![Fig. 3.5 – Maximum of a function with a softening of the nonnegativity constraint]

An example may help classify the determination of the value of the multiplier $\mu$ in the optimal solution. Let

\[
(3.21) \quad f(x; s) = a - (x + s)^2 \quad \text{with} \quad s = (-1, 0, +1)
\]

The first order conditions (FOC) for an unconstrained maximum is $f'(x; s) = -2(x + s) = 0$ with solution

\[
(3.22) \quad x^* = \begin{cases} 
+1 & \text{for } s = +1 \\
0 & \text{for } s = 0 \\
-1 & \text{for } s = -1 
\end{cases}
\]

Introducing the nonnegativity constraint on $x$, the Lagrangean of the maximization problem is

\[
(3.23) \quad L(x, \mu; s) = a - (x + s)^2 - \mu(-x)
\]

The FOC are

\[
\frac{dL(\cdot)}{dx} = -2(x^* + s) + \mu^* = 0
\]

\[
\frac{dL(\cdot)}{d\mu} = x^* \geq 0 \quad \text{with the added slack condition} \quad \mu^*x^* = 0
\]

The optimal solution is
A final remark. The Kuhn-Tucker condition is actually a necessary and sufficient condition for a maximum. If the point \( x^* = 0 \) happened to be a minimizer rather than a maximizer of the function \( f(x) \) in the feasible region, movement away from the boundary into the strictly positive part of the feasible region would be associated with a positive value of the derivative \( f'(x) \), as depicted in Fig. 3.6. But this would violate the condition in the first line of (3.17).

\[
\begin{cases}
  +1,0 \\
  0,0 \\
  -1,2
\end{cases}
\quad \text{for} \quad s = \begin{cases}
  -1 \\
  0 \\
  +1
\end{cases}
\]  

Fig. 3.6 – The Kuhn-Tucker condition excludes a maximum on the boundary if \( f'(0) > 0 \)

3.4 Constrained optimization with wealth and nonnegativity constraints

Putting together the results of the preceding Sections 3.2 and 3.3 we are now ready to address the full maximization problem initially stated and here repeated

\[
\begin{align*}
\max & \quad u(x) \\
\text{subject to} & \quad p \cdot x - w \leq 0 \text{ and } x \geq 0
\end{align*}
\]  

with, as usual, \( u(x) \) a continuous twice differentiable utility function, \( p \not\perp 0 \) and \( w > 0 \). The Lagrangean associated with this maximization problem is

\[
L(x, \lambda, \mu) = u(x) - \lambda (p \cdot x - w) - \mu \cdot (-x)
\]
Taking into account the fact that the wealth constraint is necessarily satisfied with the equal sign, the critical values are determined by the following conditions

$$\frac{\partial L}{\partial x} = \nabla x L = \nabla u(x^*) - \lambda^* p + \mu^* = 0$$

(3.28) $$\frac{\partial L}{\partial \lambda} = p \cdot x^* - w = 0$$

$$\frac{\partial L}{\partial \mu} = \nabla \mu L = x^* \geq 0$$ with the slack condition $\mu^*_i = 0$ if $x^*_i > 0, \mu^*_i > 0$ if $x^*_i = 0$

Using the Kuhn-Tucker conditions, the solution can be written in the compact form

$$\nabla u(x^*) - \lambda^* p \leq 0$$

(3.29) $$x^* \left[ \nabla u(x^*) - \lambda^* p \right] = 0$$

$$p \cdot x^* - w = 0$$

It is worth looking with a little attention into the meaning of the multipliers of the nonnegativity constrains, which can be retrieved from the slack conditions $\mu^*_i \cdot x^*_i = 0$, adapting to the context of the utility maximization problem the considerations made earlier. If we have an interior optimum, i.e. if the variables are all strictly positive, the $\mu^*_i$'s are all zero. If we are, on the contrary, at a boundary solution, some $\mu^*_i$'s may be equal to zero, but at least one must be positive.

Consider the two commodity case and the boundary solution depicted by point $B$ in Fig. 3.7, where $x^*_1 > 0$ and $x^*_2 = 0$. Note first that the tangency condition is not met, in fact, as suggested by the diagram, the marginal rate of substitution is less than the price ratio. Note second, concerning the role of the nonnegativity constraint, that if the constraint could be relaxed to, say, $x^*_2 \geq -\overline{x}_2$, the optimal solution would occur on a higher indifference curve at point $B'$. A positive value of the multiplier signals the gain in utility obtainable from a softening of the constraint. Note that the opposite would take place if we were to consider a tightening of the constraint to $x^*_2 \geq \gamma_2$ where $\gamma_2 > 0$ could represent a subsistence level, that must be respected in the consumption of commodity 2. Utility maximization would in this case occur on a lower indifference curve at point $B''$ in Fig. 3.7.

---

15 The Stone-Geary utility function reflects precisely this type of assumption and gives rise to the Linear Expenditure System of demand functions (see R. Stone, 1954). The demand functions of Stone’s model are considered in Lecture Note 4.
Fig. 3.7 – Effect on maximum utility of removing (or tightening) a bounding non negativity constraint on the consumption of commodity 2

3.5. The expenditure minimization problem (EMP)

The expenditure minimization problem will be examined in Lecture Note 5. It is briefly considered here only for the aspect concerning the formulation of the Lagrange’s function. The is

\[
\min \ p \cdot x \\
\text{such that } u(x) \geq u \text{ and } x \geq 0
\]  

(3.30)

for given prices \( p \not\leq 0 \) and a given predefined, required level of utility that must be reached. Since \( p \cdot x \) is continuous and the constraint set is non empty, convex, closed and bounded from below, the EMP has solution.

In setting up the Lagrangean for this problem we again aim at a solution in which the sign of all the multipliers should correctly indicate the direction of change of the value function from a small tightening of the constraints, i.e. from an increase in the prescribed level of utility to be reached and from the possibility that a strictly positive consumption of some commodities be required. To this end, we follow the convention that all the constraint must be set up in the “greater than or equal” form and the Lagrangean multipliers preceded by a minus sign.
The Lagrangean of the problem is therefore

\[ L(x, \lambda, \mu) = p \cdot x - \lambda (u(x) - u) - \mu \cdot (x) \]  

(3.31)

Taking account of the fact that the utility constraint is necessarily satisfied with the equal sign, the critical values are determined by the following conditions

\[ \frac{\partial L}{\partial x} = \nabla_x L = p - \lambda \nabla u(x) - \mu = 0 \]

(3.32)

\[ \frac{\partial L}{\partial \lambda} = u(x) - u = 0 \]

\[ \frac{\partial L}{\partial \mu} = \nabla_\mu L = x \geq 0 \text{ with the slack condition } \mu^*_i = 0 \text{ if } x^*_i > 0, \mu^*_i > 0 \text{ if } x^*_i = 0 \]

Using the Kuhn-Tucker conditions, the solution can be written in the compact form

\[ p - \lambda^* \nabla u(x^*) \geq 0 \]

(3.33)

\[ x^* \left[ p - \lambda^* \nabla u(x^*) \right] = 0 \]

\[ u(x^*) - u = 0 \]

The multiplier \( \lambda \) indicates the impact on the minimum expenditure when the constraint is tightened, i.e. when the prescribed level of utility \( u \) is increased: the consequence is an increase in the level of expenditure, \( \lambda^* \) is, therefore, positive. Analogously, as shown in Fig. 3.8, the expenditure required to attain a given level of utility is increased if the nonnegative constraint on the consumption of commodity 2 is made more stringent with \( x_2 \geq \gamma_2 \). A clear duality relation between the UMP and the EMP appears from this point of view: the signs of the multipliers are the same, but in the UMP they reflect the assumption of a softening of the constraints, whereas in the EMP they reflect their tightening.

---

16 Note that we can arrive at the formulation of equation (3.31) by turning the minimization into a maximization problem. We can in fact set up the problem in terms of the maximization of \(-p \cdot x\) subject to the minimum utility constraint and to the non negativity constraints. Following the rules given in the previous paragraph, the Lagrangean would be \( L(x, \lambda, \mu) = -p \cdot x - \lambda (u(x) - u) - \mu \cdot (-x) \). Multiplying by \(-1\) we obtain (3.31).

17 We maintain, for a reason of convenience of reference, the denomination of nonnegativity constraint also in this case, although it would be more proper to speak of a minimum positive binding constraint.
3.6 The role of concavity and quasiconcavity in optimization problems

The objective of this final Section is to state the relation between the second order conditions for a maximum or a minimum which, as examined in Section 3.2 and in Appendix A.3.2, take the form of properties of appropriate Hessian or Bordered Hessian matrices, with the notions of concavity (convexity) and quasiconcavity (quasiconvexity) defined in Section 2.2 of Lecture Note 2 precisely in terms of those matrices.

Let us first consider an unconstrained optimization problem. Given a generic twice differentiable function \( f(x) \), assume that the vector \( x^* \), not necessarily unique, solves the first order conditions \( \nabla f(x^*) = 0 \) and determines, therefore, the critical points which may be maxima or minima or saddle points. The second order condition distinguishes among these as follows

\[
(3.34) \quad x^* \text{ is } \begin{cases} 
\text{a maximum} & \text{if } H(x^*) \text{ is negative semidefinite} \\
\text{a minimum} & \text{if } H(x^*) \text{ is positive semidefinite} \\
\text{a saddle point} & \text{indefinite}
\end{cases}
\]

where \( H(x^*) \) is the Hessian matrix of the function \( f(x) \) evaluated at \( x^* \). Taking account of the properties of concavity and convexity of the function \( f(x) \), we can, therefore, conclude with the following proposition.
**Proposition 3.1** Let $f(x)$ be a twice continuously differentiable function on $D \subseteq \mathbb{R}^L$, then $x^*$ in the interior of $D$

(i) is a maximum if $\nabla f(x^*) = 0$ and $f(x)$ is concave;

(ii) is a minimum if $\nabla f(x^*) = 0$ and $f(x)$ is convex.

Furthermore:

(i) if $\nabla f(x^*) = 0$ and $f(x)$ is strictly concave, then $x^*$ is a maximum of $f(x)$;

(ii) if $\nabla f(x^*) = 0$ and $f(x)$ is strictly convex, then $x^*$ is a minimum of $f(x)$.

It is, incidentally, worth noting that Definition 2.9 of Lecture Note 2 of a concave function has an important implication for the identification of the properties of the solutions of a maximization problem. We state this as a corollary.

**Corollary 3.1.** If $f(x)$ is a twice continuously differentiable concave function on $D \subseteq \mathbb{R}^L$ and $x^*$ in the interior of $D$ is a local maximum, then for $y \neq x^* \in \text{int} D$

$$\nabla f(x^*)(y-x^*) \leq 0 \quad \text{implies} \quad f(y) \leq f(x^*)$$

If (3.35) is verified for all $y \in \text{int} D$, then $x^*$ is a global, not necessarily unique, maximizer of $f(x)$.

A similar Corollary follows from the definition of a twice continuously differentiable convex function simply turning the sign \(\leq\) into the sign \(\geq\).

Let us turn to a constrained optimization problem. The second order condition involves now the properties of the Hessian matrix on the subspace $Z \subseteq \mathbb{R}^L_{++}$ determined by the constraints imposed to the optimization problem. The second order condition distinguishing among maxima, minima and saddle points is as follows

$$x^* \quad \text{is} \quad \begin{cases} \text{a maximum} & \text{if } H(x^*) \text{ is negative semidefinite in the subspace } Z \subseteq \mathbb{R}^L_{++} \\ \text{a minimum} & \text{if } H(x^*) \text{ is positive semidefinite in the subspace } Z \subseteq \mathbb{R}^L_{++} \\ \text{a saddle point} & \text{indefinite} \end{cases}$$

Connecting finally the second order conditions just stated to quasiconcave and quasiconvex functions, we may conclude with the following proposition.

---

18 Actually, the result holds for a $C^1$. 

Proposition 3.2 Let $f(x)$ be a twice continuously differentiable function on $D \subseteq \mathbb{R}^n$, then $x^*$ in the interior of $D$

(i) is a maximum if $\nabla f(x^*) = 0$ and $f(x)$ is quasiconcave;
(ii) is a minimum if $\nabla f(x^*) = 0$ and $f(x)$ is quasiconvex.

Furthermore:

(i) if $\nabla f(x^*) = 0$ and $f(x)$ is strictly quasiconcave, then $x^*$ is a maximum of $f(x)$;
(ii) if $\nabla f(x^*) = 0$ and $f(x)$ is strictly quasiconvex, then $x^*$ is a minimum of $f(x)$. 
Appendix A.3 The analytics of the utility maximization problem

Appendix 3.A.1 Unconstrained optimization. At the end of Section 3.2 an intuitive explanation, based on economic considerations, was given regarding the role of the first and second order necessary condition for the determination of a maximizing solution. As anticipated, we provide here an analytical explanation with reference to the very simple one variable no constraint case.\(^{19}\). In order to clarify the distinction between second order necessary and sufficient conditions for a maximum, we consider

Let \( f(x); x \rightarrow \mathbb{R} \) be a twice continuously differentiable function of the scalar variable \( x \).

Suppose that \( x^* \) is a local maximizer of \( f(x) \); suppose in other words that there exists an open neighborhood of \( x^* \), say \( A \subset \mathbb{R} \), such that for all small arbitrary deviations \( z = x - x^* \in \mathbb{R} \) about \( x^* \) in any direction, we have

\[
(A3.1) \quad f(x^*) \geq f(x^* + z)
\]

Expanding the function on the right hand side of (A3.1) in a second order Taylor’s series about the point \( x^* \) (where \( z = 0 \)) and disregarding higher order terms, which for \( z \) small will be dominated by the quadratic term, we have

\[
(A3.2) \quad f(x^* + z) = f(x^*) + \frac{df}{dx}(x^*)z + \frac{1}{2} \frac{d^2f}{dx^2}(x^* + \theta z)z^2
\]

with \( 0 < \theta < 1 \). Substituting in (A3.1) we obtain the inequality

\[
(A3.3) \quad \frac{df}{dx}(x^*)z + \frac{1}{2} \frac{d^2f}{dx^2}(x^* + \theta z)z^2 \leq 0
\]

which must hold for any arbitrary small variation \( z \).

By dividing both sides by \( z \) and taking the limit as \( z \) approaches zero, we have

\[
(A3.4) \quad \begin{cases} \frac{df}{dx}(x^*) \leq 0 & \text{if} \quad z > 0 \\ \frac{df}{dx}(x^*) \geq 0 & \text{if} \quad z < 0 \end{cases}
\]

\(^{19}\) We follow here Intrilligator’s (2000), pp. 22–25) presentation and reproduce his illuminating diagram.
Since the fundamental inequality (A.3.3) must be satisfied for all small arbitrary deviations \( z \), we must have, as a first order necessary condition, that the first derivative of the function be equal to zero at the local maximum

\[
(A3.5) \quad \frac{df}{dx}(x^*) = 0
\]

Turning to the second order condition, we can first argue by exclusion. Using the first order condition (A3.5), we can rewrite (A3.2) as

\[
(A3.6) \quad f(x^* + z) - f(x^*) = \frac{1}{2} \frac{d^2 f}{dx^2}(x^* + \theta z)z^2
\]

Since \( z^2 \) is always positive, it follows that if the second order derivative \( \frac{d^2 f}{dx^2}(x^* + \theta z) \) is positive, there exists \( z \), in the small neighborhood \( A \subset \mathbb{R} \) of \( x^* \), for which \( f(x^* + z) - f(x^*) > 0 \), contradicting the assumption that \( x^* \) is a maximizer. We conclude that for \( x^* \) to be a local maximizer of \( f(x) \) the second order derivative must be negative or zero, namely

\[
(A3.7) \quad \frac{d^2 f}{dx^2}(x^*) \leq 0
\]

Conditions (A3.5) and (A3.7) are then the first and second order necessary conditions for a local maximum at \( x^* \).

The extension of the unconstrained optimization to the case of a vector variable \( x \in \mathbb{R}^L \) can be analyzed with a similar procedure. Suppose that \( x^* \) is a local maximizer of \( f(x) \) and let \( z \in \mathbb{R}^L \) be a small arbitrary deviation about \( x^* \). We have

\[
(A3.8) \quad f(x^*) \geq f(x^* + hz)
\]

with \( h \) an arbitrary small positive number. Expanding the right hand side of (A.3.9) in a Taylor series about the point \( h = 0 \), we obtain

\[
(A3.9) \quad f(x^* + hz) = f(x^*) + h \nabla f(x^*) \cdot z + \frac{1}{2} h^2 \cdot Hf(x^* + hz)hz
\]

---

20 The same result could have been reached using in (A3.2) the first order condition and taking account of the fact that \( z^2 \) is always positive. The same approach is used by MWG(p. 933) to show that a function of a vector variable is concave if and only if the Hessian matrix is negative semidefinite.
where $\nabla f(x^*)$ is the vector of partial derivatives of the function, which must be equal to zero for a critical point of the function, and $Hf(x^* + \theta hz)$ the Hessian matrix of $f(x)$, which must the negative semidefinite for a maximum.

**Sufficient conditions** for a strict local maximum at $x^*$ are, in the scalar variable case, that the first derivative vanishes and, in particular, that the second order derivative be negative for all $x \neq x^*$. The second order sufficient condition for a maximum is therefore

$$d^2f(x^*) < 0$$

Give the stationarity condition (A3.5), changing the sign in relations (A3.7) and (A3.8) from less than or equal to greater than or equal and from less than to greater then we obtain the second order necessary and sufficient conditions of a local minimum.

Fig. A3.1 in which the diagrams of the three functions: $f(x)$, $\frac{df}{dx}(x)$ and $\frac{d^2f}{dx^2}(x)$ are depicted one below the other, illustrates the difference between necessary and sufficient second order conditions:

the points $x^*$ and $x^{****}$ are both strict local maxima: the first derivative vanishes and the second derivative is negative;

the point $x^{**}$ is a strict minimum: the first derivative vanishes and the second derivative is positive;

the inflexion point $x^{***}$ is the crucial one for our purpose: the first order condition of zero slope is met, but the second order derivative is also equal to zero, but changing sign from negative to positive as we depart from $x^{***}$. The sufficient condition for a local maximum is not met. This shows that the first and second order necessary conditions for a maximum are not sufficient; they, in effect, coincide with the analogous first and second order conditions for a minimum. A horizontal inflection point as $x^{***}$ sign is neither a local maximum, nor a local minimum.

---

21 In the vector variable case the second order sufficient condition is that Hessian be negative definite.
Appendix 3.A.2 Constrained optimization: Lagrange’s theorem and function

A3.2.1 Lagrange’s Theorem

Generalizing the presentation in Section 3.2 of the utility maximization problem subject to the wealth constraint, consider now the problem of maximizing an objective function \( f(x) \) under one equality constraint which we assume to be convex and write in the more general, not necessarily linear, form \( g(x) = 0 \), i.e. the problem\(^{22}\)

\[
\max_{x \in \mathbb{R}} f(x) \quad \text{s.t.} \quad g(x) = 0 
\]

\(^{22}\) We consider only the maximization problem since the equality constrained minimization problem can be handled by maximizing the negative of the objective function subject to the same constraints.
Following the approach of Section A3.1, let $f(x) : \mathbb{R}^L \to \mathbb{R}$ be a twice continuously differentiable function of the vector variable $x$. Suppose that $x^*$ is a local maximizer of $f(x)$ subject to the twice differentiable and convex constraint $g(x) = 0$; suppose in other words that there exists an open neighborhood of $x^*$, say $A \subset \mathbb{R}^L$, such that for all small arbitrary deviations $z = x - x^* \in \mathbb{R}^L$ about $x^*$ in any direction, we have

\[(A3.12) \quad f(x^*) \geq f(x^* + hz)\]

with $h$ an arbitrary small positive number. Expanding the right hand side of (A.3.12) in a Taylor series about the point $h=0$, we obtain

\[(A3.13) \quad f(x^* + hz) = f(x^*) + h \nabla f(x^*) \cdot z + \frac{1}{2} hz \cdot Hf(x^* + \theta hz) hz\]

where $\nabla f(x^*)$ is the vector of partial derivatives, which must be equal to zero for a critical point of the function, and $z \cdot Hf(x^*) z$ is a quadratic form, which must be negative semidefinite (negative definite) for a maximizing solution.

As we have seen in Lecture Note 2, Section 2.2, the presence of a constraint requires that the properties of the Hessian $Hf(x)$ be evaluated on the subspace $Z = \{ z \in \mathbb{R}^L \mid \nabla f(x^*) \cdot z = 0 \}$.

We have shown in Section 3.2 above that this property is verified in the utility maximization problem by a particular Bordered Hessian, namely by the Hessian of $u(x)$ bordered by the vector of commodity prices, which define the slope of the tangent line to the indifference curve at $x^*$. In order to better understand the meaning of these particular bordering elements it is useful to take into consideration a somewhat more general constrained maximization problem. We suppose, in fact, that the constraint need not be necessarily linear, as the wealth constraint is, but that it may be strictly convex.

The plan of this Section is to state first, without proof, Lagrange’s theorem: We then show: (i) that the critical values of the Lagrangian function $(x^*, \lambda^*)$ do satisfy the conditions for $x^*$ to be a maximizer of $f(x)$ subject to the constraint $g(x) = 0$; and (ii) that the Bordered Hessian $\tilde{H}^b(x^*)$, defined in (3.13), of the utility maximization problem under the wealth constraint is just a particular case of the more general problem of maximization under a convex constraint.
We will go back later to the linear form of the wealth constraint to establish the connection with the presentation in Section 3.2; we maintain the assumption of the existence of an internal solution and disregard the nonnegativity constraints on the variables. The approach may be extended to the consideration of a set of $M$ constraints with $L \geq M$ for, otherwise, the constraint set will in general be empty.\footnote{For the more general approach, in particular for the case of two constraints and the relative graphical presentations see MWG (pp. 956-9963), JR (pp. 577-601), Intrilligator (pp. 28-38).}

The feasible point $x^* \in C$ is a local maximizer of $f(x)$ if there exists an open neighborhood of $x^*$, say $A \subseteq \mathbb{L}$, such that $f(x^*) \geq f(x)$ for all $x \in A \cap C$; $x^*$ is a global maximizer of $f(x)$ if $f(x^*) \geq f(x)$ for all $x \in \mathbb{L}$.

**Lagrange’s Theorem.** Let: (i) the objective function $f(x)$ and the constraint function $g(x)$ be twice continuously differentiable and increasing in $x$; (ii) $x^*$ be an interior solution of the optimization problem \( \text{A3.11} \); (iii) the vector of partial derivatives of the constraints is positive, $\nabla f(x) > 0$. Then there is a positive number $\lambda^* \in \mathbb{R}$, the Lagrangean multiplier of the constraint, such that

\[
\frac{\partial f(x^*)}{\partial x_l} = \lambda^* \frac{\partial g(x^*)}{\partial x_l} \quad l = 1, \ldots, L
\]

or in vector form notation

\[
\nabla f(x^*) = \lambda^* \nabla g(x^*)
\]

In words, the gradient vector of the objective function is proportional to a linear combination of the gradient vectors of the constraint function.\footnote{The extension to the case of $M > 1$ constraints requires to introduce the constraint qualification, namely the condition that the Jacobian matrix $M \times N$ be of rank $M$. In this case, Lagrange’s theorem asserts the existence of a vector of $M$ positive multipliers, one for each constraint, such that the gradient vector of the objective function is equal to a linear combination of the gradient vectors of the constraint functions.}

We refer for a proof of Lagrange’s theorem to the references indicated in footnote 23 and pass directly to an examination of Lagrange’s function and its properties.

The Lagrange’s function associated to the maximization problem \( \text{A.3.11} \) is

\[
L(x, \lambda) = f(x) - \lambda g(x)
\]
The critical values \((x^*, \lambda^*)\) are the solutions of the \(L+1\) first order conditions

\[
\Delta L_x (x, \lambda) = \nabla f \left( x^* \right) - \lambda^* \nabla g \left( x^* \right) = 0 \\
\Delta L_\lambda (x, \lambda) = g \left( x^* \right) = 0
\]

(A3.17)

The \(L\) equations of the first line of (A3.17) exactly coincide with (A3.15). The fundamental result of Lagrange’s Theorem is immediately verified by the first order conditions of the associated Lagrangian function.

There remains to ascertain that the critical values \(x^*\) of the Lagrangian function do maximize the objective function \(f(x)\) subject to the constraint.

At the critical values \((x^*, \lambda^*)\) the Lagrangian function \(L(x, \lambda)\) must be stationary; this implies that the total differential \(dL\) must be equal to zero

\[
dL(x, \lambda) = \sum_{i=1}^{L} \frac{\partial f \left( x^* \right)}{\partial x_i} dx_i - g \left( x^* \right) d\lambda^* - \lambda^* \sum_{i=1}^{L} \frac{\partial g \left( x^* \right)}{\partial x_i} dx_i = 0
\]

(A3.18)

Note first that the second term on the right hand side of (A3.18) vanishes because the critical values \(x^*\) satisfy the equality constraint - conditions (A3.17). Observe next that also the third term on the right hand side of (A3.18) is equal to zero since the admissible deviations \(dx_i\) in the neighborhood of \(x^*\) must satisfy the constraint, hence \(\sum_{i=1}^{L} \frac{\partial g \left( x^* \right)}{\partial x_i} dx_i = 0\). (A3.18) thus reduces to

\[
dL(x, \lambda) = \sum_{i=1}^{L} \frac{\partial f \left( x^* \right)}{\partial x_i} dx_i = 0
\]

(A3.19)

which insures that any arbitrary small deviation from the optimal solution \(x^*\) does not increase the value of the objective function.

We can now say a bit more about the nature of the critical values \((x^*, \lambda^*)\) of the Lagrange’s function. We can say that the maximum over the choice variable \(x\) of the objective function \(f(x)\) subject to the equality constraint \(g(x) = 0\) coincides with the unconstrained maximum

\[\text{25} \text{ Going back to Section 3.2.2, in which the two variable one equality constraint case was studied, Fig. 3.3 shows that the admissible deviations } dx_i \text{ must lie on the budget constraint.}\]
over $x$ of the associated Lagrangean. We defer for the moment discussing the nature of the Lagrangean multipliers $\lambda^*$. 

**A3.2.3 Second order necessary conditions for a maximum**

The study of second order necessary and sufficient conditions, already discussed in Section 3.2 with reference to the problem of utility maximization subject to the wealth constraint, can now be extended to the case here considered of a twice differentiable equality constraint, obviously maintaining the assumption that the objective function is quasiconcave and twice differentiable. We now show that the second order condition derived using the substitution approach to the constrained maximization of the utility function coincides with the Bordered Hessian of the Lagrange’s function.

Consider the two-commodity maximization problem

$$\begin{align*}
\max_{x \in \mathbb{R}^2} f(x_1, x_2) \\
\text{s.t. } g(x_1, x_2) = 0
\end{align*}$$

(A3.20)

Following the “substitution” approach to the solution of the maximization problem, think of the constraint as defining $x_2$ in terms of $x_1$, namely $x_2(x_1)$. The slope of the constraint set is

$$\frac{dx_2}{dx_1} = -\frac{g_1}{g_2}$$

(A3.21)

Substituting $x_2(x_1)$ in the objective function, let

$$h(x_1) = f(x_1, x_2(x_1))$$

(A3.22)

be the value of the objective function subject to the constraint. Differentiating $h(x_1)$ with respect to $x_1$ and taking first and second order derivatives as we have done in Section 3.2, we respectively have

$$\frac{dh}{dx_1} = h_1 - h_2 \frac{g_1}{g_2}$$

(A3.23)

$$\frac{d^2h}{dx_1^2} = f_{11} + f_{12} \left(-\frac{g_1}{g_2}\right) + \left[f_{21} + f_{22} \left(-\frac{g_1}{g_2}\right)\right] \frac{g_1}{g_2} -$$

$$- \frac{f_2}{g_2} \left(g_1 + g_2 \left(-\frac{g_1}{g_2}\right)\right) - g_1 \left(g_{21} + g_{22} \left(-\frac{g_1}{g_2}\right)\right)$$

(A3.24)
The joint presence of the second order derivatives of the objective function and of the constraint function does not permit of a direct interpretation of this expression, which however becomes possible in connection with the introduction of the Lagrangean function (A3.16). Given the first order conditions

\[
L_1(x^*, \lambda^*) = f_1(x^*) - \lambda^* g_1(x^*) = 0 \\
L_2(x^*, \lambda^*) = f_2(x^*) - \lambda^* g_2(x^*) = 0
\]

we can define the second order partial derivatives of \( L(x, \lambda) \) at \( (x^*, \lambda^*) \)

\[
L_{11}(x^*, \lambda^*) = f_{11}(x^*) - \lambda^* g_{11}(x^*) \\
L_{12}(x^*, \lambda^*) = L_{21}(x^*, \lambda^*) = f_{12}(x^*) - \lambda^* g_{12}(x^*) \\
L_{22}(x^*, \lambda^*) = f_{22}(x^*) - \lambda^* g_{22}(x^*)
\]

Substituting in (A3.24) first equations (A3.25) and subsequently the second order derivatives (A3.26), we obtain

\[
\frac{d^2 h}{dx_i^2} = \frac{1}{g_2} \left\{ (f_{11} + \lambda^* g_{11}) g_2^2 - 2 \left( (f_{12} + \lambda^* g_{12}) g_1 g_2 + (f_{22} + \lambda^* g_{22}) g_1^2 \right) \right\} = \\
- \frac{1}{g_2} \left\{ L_{11} g_2^2 + 2 L_{12} g_1 g_2 - L_{22} g_1^2 \right\}
\]

The term in curly brackets is equal to the determinant of the matrix

\[
\text{det} \bar{H}L(x^*, \lambda^*) = \begin{vmatrix} L_{11} & L_{12} & g_1 \\ L_{21} & L_{22} & g_2 \\ g_1 & g_2 & 0 \end{vmatrix}
\]

which is the Bordered Hessian of the Lagrangean function, namely the matrix of the second order partial derivatives of \( L(x, \lambda) \) bordered by the first derivatives of the constraint, all evaluated at the critical values \( (x^*, \lambda^*) \). We can conclude that the second order necessary (sufficient) condition for a maximum is satisfied if the determinant of the bordered Hessian \( \bar{H}L(x^*, \lambda^*) \) is nonnegative (positive) for all \( x \neq x^* \).

For the interpretation of this second order condition we can look at Figg. A3.2(a) and (b). The diagrams suggest that the critical value \( x^* \) is a maximer if for any arbitrary deviation about \( x^* \) the curvature of the contour set of the objective function is greater in absolute value than the curvature of the constraint. The condition is certainly satisfied, given the assumption that the
objective function is quasiconcave, if the constraint function is convex (Fig. A3.2(a)) and *a fortiori* if it is strictly convex (Fig. A3.2(b). A failure of the second order necessary condition for a maximum would occur if this condition were not respected. Fig. A3.3 depicts such a situation: \( x^* \) cannot be a maximizer since displacement about \( x^* \) would make it possible to reach a higher contour set of the objective function.

Fig. A3.2 – Panel (a) \( g(x) = 0 \) convex; Panel (b) \( g(x) = 0 \) strictly convex

Fig. A3.3 - \( g(x) = 0 \) strictly concave

The generalization to any number of constraints \( m=1,...,M < L \) would require to consider the Bordered Hessian \( \bar{HL}(x^*, \lambda^*) \), defined in block form in (A3.29), in which \( L_{\bar{H}}(x^*, \lambda^*) \) is the Hessian matrix of the Lagrangean function and \( M(x^*, \lambda^*) \) the Jacobian matrix \( M \times L \) of the vectors of the partial derivatives of the \( M \) constraints.
and ascertain the conditions for it to be positive semidefinite (definite).\textsuperscript{26}

\textbf{Appendix A.3.3 Nonnegativity constrained optimization: Kuhn-Tucker conditions}

We consider now the case of nonnegativity constraints with reference first to the case in which the function $f(x)$ of the scalar variable $x$ is twice continuously differentiable. The approach is similar to that followed in Section A3.1: we derive conditions on the first and second order derivatives for $x^*$ to be a maximizing solution taking into account the implication of the nonnegativity constraint on $x$ for the admissible deviations. The resulting nonlinear programming problem\textsuperscript{27} is

\begin{equation}
\text{max } f(x) \\
\text{s.t. } x \geq 0
\end{equation}

Let $f(x): \mathbb{R} \rightarrow \mathbb{R}$ be a twice continuously differentiable function of the scalar variable $x$. Suppose that $x^*$ is a local maximizer of $f(x)$ subject to the constraint $x \geq 0$; suppose in other words that there exists an open neighborhood of $x^*$, say $A \subset \mathbb{R}$, such that for all small arbitrary deviations $z = x - x^* \in \mathbb{R}_+$ about $x^*$, we have

\begin{equation}
f(x^*) \geq f(x^* + hz)
\end{equation}

$h$ an arbitrary small positive number. Expanding the function on the right hand side of (A3.31) in a second order Taylor’s series about the point $x^*$ (where $z = 0$) and disregarding higher order terms, which for $z$ small will be dominated by the quadratic term, we have

\begin{equation}
f(x^* + z) = f(x^*) + \frac{df}{dx}(x^*)hz + \frac{1}{2} \frac{d^2f}{dx^2}(x^* + \theta hz)(hz)^2
\end{equation}

with $0 < \theta < 1$. Substituting in (A3.31) we obtain the inequality

\textsuperscript{26} See Lecture Note 2 for the definitions of these conditions.
\textsuperscript{27} This is the standard terminology for an optimization problem with inequality constraints.
\[ \frac{df}{dx}(x^*)hz + \frac{1}{2} \frac{d^2f}{dx^2}(x^* + \theta hz)(hz)^2 \leq 0 \]

which must hold for any nonnegative small variation \( z \).

If \( x^* \) is an interior solution \( (x^* > 0) \), all arbitrary variations of \( z \) are admissible, inasmuch as they would not lead to a violation of the nonnegativity constraint. This case coincides, therefore, with the situations of unconstrained maximization examined in Section A3.1, namely the first derivative must be equal to zero and the second derivative must be less than or equal to zero. If \( x^* \) is a boundary solution \( (x^* = 0) \), the only admissible deviations is \( z > 0 \). We would then be in the situation described by the first line of (A3.4) and conclude from the fundamental equation (A3.33), as we have done in Section A3.1, that maximization requires

\[ \frac{df}{dx}(x^*) \leq 0 \]

As explained in Section 3.3, the condition for the maximizing solution can then be written in compact form using the Kuhn-Tucker condition, both necessary and sufficient, here repeated

\[ f(x^*) \leq 0 \quad \text{with equality if } x^* > 0 \]

\[ x^* f(x^*) = 0 \]

The extension of the nonnegativity constrained optimization to the case of a vector variable \( x \in \mathbb{R}^L \) is straightforward,

**Appendix A.3.4 Inequality constrained optimization: the saddle point solution**

We finally turn to the study of the problem of maximizing an objective function subject to a set of inequality rather than to equality constraints, as well as to the condition of nonnegativity of the choice variables. We have already examined a problem of this type on Section 3.4, namely the problem of maximizing a utility function subject to a single equality constraint (the budget set) and to the condition that the quantities demanded be nonnegative. We now consider, with reference to the utility maximization problem, the following more general nonlinear programming problem

\[
\max_{x \in \mathbb{R}^L} u(x) \\
\text{s.t. } p \cdot x - w \leq 0 \\
x \geq 0
\]

(A3.36)
The formulation of the wealth constraint in the “less than or equal” form makes it possible to clarify the nature of the solution.

A convenient approach to the solution of the problem is to convert it into a maximization problem in which the only constraints are the nonnegativity constraints. We can do it by introducing a “slack variable” \( s \) so as to transform the wealth inequality wealth constraint into an equality constraint. The optimization problem becomes

\[
\max_{x \in \mathbb{R}} u(x)
\]

subject to

\[
p \cdot x + s - w = 0
\]

\[
x \geq 0, \quad s \geq 0
\]

Taking account of the nonnegativity constraints using the Kuhn-Tucker conditions as we have done in Section 3.4, relation (3.29), we consider the Lagrangean function

\[
L(x, \lambda, s) = u(x) - \lambda (p \cdot x + s - w)
\]

The critical values are determined by the following set of conditions

\[
\frac{\partial L}{\partial x} = \nabla u(x^*) - \lambda^* p \leq 0 \quad \text{if} \quad x^* > 0
\]

\[
x^* \cdot [\nabla u(x^*) - \lambda^* p] = 0
\]

\[
x^* \geq 0
\]

\[
\frac{\partial L}{\partial \lambda} = p \cdot x^* + s^* - w = 0
\]

\[
\frac{\partial L}{\partial s} = -\lambda^* \leq 0
\]

\[
s^* \lambda^* = 0
\]

\[
s^* \geq 0
\]

The first three lines correspond to (3.29) above; the forth line is the derivative of the Lagrangean with respect to the multiplier \( \lambda \), with the wealth constraint formulated as an equality. The last three lines are Kuhn-Tucker conditions with reference to the slack variable \( s \). Note in particular that the fifth line implies that in the optimal solution the Lagrangean multiplier must be weakly positive.

Eliminating the slack variable \( s \), that was introduced only in order to convert the general nonlinear programming problem into a nonlinear programming problem with only nonnegativity constraints of the variable, the maximization problem becomes
\[
\max_{x \in \mathbb{R}} u(x) \\
\text{s.t. } p \cdot x - w \leq 0 \\
x \geq 0
\]

(A3.40)

The Kuhn-Tucker conditions for the determination of the critical values of the Lagrangean function

(A3.41) \[ L(x, \lambda) = u(x) - \lambda (p \cdot x - w) \]

are

\[
\frac{\partial L(\cdot)}{\partial x} = \nabla u(x^*) - \lambda^* p \leq 0 \quad \text{if } x^* > 0 \\
x^* \left[ \nabla u(x^*) - \lambda^* p \right] = 0 \\
x^* \geq 0
\]

(A3.42)

\[
\frac{\partial L(\cdot)}{\partial \lambda} = p \cdot x^* - w \leq 0 \\
\lambda^* \left[ p \cdot x^* - w \right] = 0 \\
\lambda^* \geq 0
\]

Observing the different direction of the inequality signs in line 1 and in line 4 of (A3.42) and recalling that maximization requires the “less than or equal” sign as in line 1, we note that the “greater than or equal” sign in line 4 reflects the first order condition of a minimization problem. We conclude that the solution \((x^*, \lambda^*)\) is a saddle point solution, maximizing in the vector of choice variables \(x\) and minimizing in the Lagrange’s multiplier \(\lambda\):

(A3.43) \[ L(x, \lambda^*) \leq L(x^*, \lambda^*) \leq L(x^*, \lambda) \]

The economic meaning of \(\lambda^*\) can be further clarified. We had named the value of the multiplier in the optimal solution of the utility maximization problem under the wealth equality constraint the marginal utility of wealth or the shadow price of wealth. We are now in a position to qualify that assertion in the sense that it is the lowest shadow price of wealth; a higher price would warrant attaining a utility level higher than \(u(x^*)\).
Generalizing to the case of a generic function \( f(x) \) and of several convex inequalities constraint \( g_m(x) \leq 0, \ m = 1,...M < L \), we can state the following Theorem, which we have proved for the case of the single wealth constraint.

**Kuhn-Tucker Theorem.** Let: (i) the objective function \( f(x) \) be quasiconcave and the constraint functions \( g_m(x) \) be convex and both be twice continuously differentiable and increasing in \( x \); (ii) the \( M \times L \) Jacobian matrix \( M(x^*, \lambda^*) \) of the vectors of the partial derivatives of the \( M \) constraints satisfy the constrain qualification. Then \( x^* \) solves the nonlinear programming problem (A3.40) if and only if \( (x^*, \lambda^*) \) is the saddle point solution of the Lagrangean function (A3.41).

### References


---

28 The constraint qualification requires that the matrix \( M(x^*, \lambda^*) \) be of rank \( M \).