Lecture 6: Optimization.

Unconstrained Optimization

1. Given a real valued function, \( y = f(x) \) we will be concerned with the existence of extreme values of the dependent variable \( y \) and the values of \( x \) which generate these extrema. (maxima or minima) The function \( f(x) \) is called the objective function and the independent variable \( x \) is called the choice variable. The problem of finding the value or the set of values of the choice variable which yield extrema of the objective function is called optimization. In order to avoid boundary optima, we will assume that \( f : X \rightarrow \mathbb{R} \), where \( X \) is an open interval of \( \mathbb{R} \). All of the optima characterized will be termed interior optima.

2. Definitions

\textit{Def 1:} \( f \) has a \textit{local maximum} at a point \( x^0 \in X \) if \( \exists N(x^0) \) such that \( f(x) - f(x^0) < 0 \) for all \( x \in N(x^0), x \neq x^0 \).

\textit{Def 2:} \( f \) has a \textit{local minimum} at a point \( x^0 \in X \) if \( \exists N(x^0) \) such that \( f(x) - f(x^0) > 0 \) for all \( x \in N(x^0), x \neq x^0 \).

\textit{Def 3:} A real valued function \( f(x), f : X \rightarrow \mathbb{R} \), has an \textit{absolute or global maximum (minimum)} at a point \( x^0 \in X \), if \( f(x) < f(x^0) \ (f(x) > f(x^0)) \), for all \( x \neq x^0 \), such that \( x \in X \).

Illustrations:

3. Optimization and the First Derivative Test

\textit{Proposition 1.} If \( f \) has a local maximum or minimum at \( x^0 \in X \), then \( f'(x^0) = 0 \).
Points at which \( f' = 0 \) are called critical values. The image \( f(x^0) \) is called a critical value of the function and \( x^0 \) is called a critical value of the independent variable. This proposition says that each maximum or minimum is a critical value of the function.

4. We would like to be able to distinguish between maxima and minima and to rule out inflection points. Thus, we must develop tools to study the curvature of the function.

a. Second and Higher Derivatives

Since the function \( f'(x) \) is itself a function of an independent variable \( x \), this fact suggests the fact that the function \( f'(x) \) may be itself a differentiable function.

**Def 1:** Given that \( y = f(x) \), the derivative with respect to \( x \) of the differentiable function \( f'(x) = \frac{dy}{dx} \) is the second derivative of \( y \) with respect to \( x \)

and is denoted by \( f''(x) \) or \( \frac{d^2y}{dx^2} \).

**Notation:** Sometimes \( \frac{d^2y}{dx^2} \) is written

\[
\frac{d^2y}{dx^2} \equiv \frac{d}{dx} \left( \frac{dy}{dx} \right) = \frac{d}{dx} f'(x) = f''(x) = \frac{d}{dx} f''(x) = f'''(x) = \frac{d}{dx} f'''(x) = f^{(2)} = f_{xx}
\]

**Examples:**

**#1** Let \( y = f(x) = x^3 + x^2 \), then \( f'(x) = 3x^2 + 2x \)

\( f^{(2)}(x) = f''(x) = 6x + 2. \)

As long as the differentiability condition is met after successive differentiation, we may define third and higher order derivatives. That is, if the function \( f''(x) \) is differentiable, then we may calculate \( f'''(x) = \frac{d}{dx} f''(x) \) for the third derivative. In general if the \( i \)th derivative function is differentiable we may calculate the \( i + 1 \) derivative, i.e.

\[
\frac{d}{dx} f^{(i)}(x) = f^{(i+1)}(x)
\]

**Example**
Let \( y = x^4 + x^3 + x \), find \( f'(x) \), \( f''(x) \), \( f'''(x) \), \( f''''(x) \).

\[
\begin{align*}
\frac{dy}{dx} &= 4x^3 + 3x^2 + 1 \\
f''(x) &= 12x^2 + 6x \\
f'''(x) &= 24x + 6 \\
f''''(x) &= 24.
\end{align*}
\]

b. Interpretation of the second derivative.

The first derivative of the function \( y = f(x) \), \( f'(x) \), measures the rate of change of the function \( f(x) \) or the rate of change of \( y = f(x) \). The second derivative of \( y = f(x) \), \( f''(x) \) measures the rate of change of the first derivative function, \( f'(x) \).

We observe the following facts:

(i) \( f'(x^0) > 0 \Rightarrow f(x) \) tends to increase

(ii) \( f'(x^0) < 0 \Rightarrow f(x) \) tends to decrease

(iii) \( f''(x^0) > 0 \Rightarrow f'(x) \) tends to increase

(iv) \( f''(x^0) < 0 \Rightarrow f'(x) \) tends to decrease

We also observe the following combination of facts.

(i) \( f'(x^0) > 0 \) and \( f''(x^0) > 0 \Rightarrow \) the slope of the curve at \( x^0 \) is positive & increasing.

(ii) \( f'(x^0) > 0 \) and \( f''(x^0) < 0 \Rightarrow \) the slope at \( x^0 \) is positive and decreasing

(iii) \( f'(x^0) < 0 \) and \( f''(x^0) > 0 \Rightarrow \) the slope at \( x^0 \) is negative and increasing.

That is \( f'(x^0) \) is negative and becoming less negative.

(iv) \( f'(x^0) < 0 \) and \( f''(x^0) < 0 \Rightarrow \) the slope at \( x^0 \) is negative and decreasing.

That is, its negative and becoming more negative.

From this analysis, we conclude that for a local maximum, it suffices that \( f'' \) be negative at the critical point and for a local minimum it suffices that \( f'' > 0 \) at the critical point.
c. Concave functions and the second derivative.

Def 1: A set $X \subseteq \mathbb{R}^N$ is convex if for any $x, x^1 \in X, \alpha \in [0, 1], \alpha x + (1-\alpha) x^1 \in X$.

Def 2: A real-valued fn $f(x), f: X \rightarrow \mathbb{R}, X \subseteq \mathbb{R}$ (X convex) is concave if for any $x, x^1 \in X, \alpha \in [0, 1], f(\alpha x + (1-\alpha) x^1) \geq \alpha f(x) + (1-\alpha) f(x^1)$.

Def 3: A real-valued fn $f(x), f: X \rightarrow \mathbb{R}, X \subseteq \mathbb{R}$ (X convex) is strictly concave if for any $x \neq x', x, x' \in X, \alpha \in (0, 1), f(\alpha x + (1-\alpha) x') > \alpha f(x) + (1-\alpha) f(x')$.

Theorem 1: If a fn $f(x), f: X \rightarrow \mathbb{R}$, such that $X$ is an open interval $X \subseteq \mathbb{R}$, is differentiable over $X$, it is strictly concave if and only if $f'(x)$ is decreasing over $x$. If $f''$ exists and is negative, for all $x$, then $f$ is strictly concave.

Remark 1: A function is convex or strictly convex if -$f$ is concave or strictly concave, respectively. Thus, the above results apply to convex functions. The strict convexity of a differentiable function is equivalent to its first derivative being an increasing function and the positivity of its second derivative implies strict convexity.

Remark 2: The definition of strict concavity for a function of $n$ variables is analogous to the above definition. In this case $f: X \rightarrow \mathbb{R}$, where $X$ is a subset of $\mathbb{R}^N$. We have

Def 4. A real-valued fn $f(x), f: X \rightarrow \mathbb{R}, X \subseteq \mathbb{R}^N$ (X convex) is strictly concave (convex) if for any $x \neq x', x, x' \in X, \alpha \in (0, 1), f(\alpha x + (1-\alpha) x') > (>) \alpha f(x) + (1-\alpha) f(x')$.

The second derivative characterization for a convex or strictly convex function is as follows:

Theorem 2. Let $f: f: X \rightarrow \mathbb{R}, X \subseteq \mathbb{R}^N$ (X convex). Suppose that $f$ is twice continuously differentiable on $X$. If $d^2f = dx'Hdx$ is negative (positive) definite for all $x \in X$, then $f$ is strictly concave (convex). If $d^2f$ is negative (positive) semidefinite for all $x \in X$, then $f$ is concave (convex).

5. A sufficiency result
Proposition 2. Let \( f \) be twice differentiable. Let there exists an \( x^o \in X \) such that \( f'(x^o) = 0 \).

(i) If \( f''(x^o) < 0 \), then \( f \) has a local maximum at \( x^o \). If, in addition, \( f'' < 0 \) for all \( x \) or if \( f \) is strictly concave, then the local maximum is a unique global maximum.

(ii) If \( f''(x^o) > 0 \), then \( f \) has a local minimum at \( x^o \). If, in addition, \( f'' > 0 \) for all \( x \) or if \( f \) is strictly convex, then the local minimum is a unique global minimum.

The zero derivative condition is called the \textit{first order condition} and the second derivative condition is called the \textit{second order condition}.

Example: #1 Let \( f(x) = x + x^{-1} \) we know that \( f'(x) = 1 - x^{-2} \) and \( x = \pm 1 \). Now calculate \( f''(x) \)

\[
f''(x) = 2x^{-3}.
\]
At \( x = 1 \), \( f''(1) = 2 > 0 \), so that 1 is a local min. At \( x = -1 \), \( f'' = -2 \), so that -1 is a local max.

#2. Let \( f = ax - bx^2 \), \( a,b > 0 \) and \( x > 0 \). Here \( f' = 0 \) implies \( x = a/2b \).

Moreover, \( f'' = -2b < 0 \), for all \( x \). Thus, we have a global max.

Existence

In the above discussion, we have assumed the existence of a maximum or a minimum. In this section, we wish to present sufficient conditions for the existence of an extremum. We have the following result.

Proposition. Let \( f : X \to \mathbb{R}, \ X \subset \mathbb{R} \), where \( X \) is a closed interval of \( \mathbb{R} \). Then if \( f \) is continuous, it has a maximum and a minimum on \( X \). If \( f \) is strictly concave, then there is a unique maximum on \( X \). If \( f \) is strictly convex, then it has a unique minimum on \( X \).

This proposition includes boundary optima in its assertion of existence. To insure existence of interior optima, one must show that boundary points are dominated by some interior point. If those boundary points are not part of the domain of the function, then one can take the limit of the function at boundary points and show that those limit points are dominated by some interior point.
An n-Variable Generalization

1. Let \( y = f(x) \), where \( x \in \mathbb{R}^n \) and \( f : X \to \mathbb{R} \), where \( X \) is an open subset of \( \mathbb{R}^n \). A set \( X \) is an open subset of \( \mathbb{R}^n \) if \( \forall \ x \in X \exists \ N(x) \subset X \). In this case \( N(x^o) \) is defined as the set of points within an \( \varepsilon \) distance of \( x^o \):

\[
N(x^o) = \{ x \mid [\sum_{i=1}^{n} (x_i - x^o_i)^2]^{1/2} < \varepsilon, \varepsilon > 0 \}.
\]

2. The definition of a local maximum or minimum is as follows:

*Def.* \( f \) has a local maximum (minimum) at a point \( x^o \in X \) if \( \exists \ N(x^o) \) such that for all \( x \in N(x^o) \), \( f(x) - f(x^o) < 0 \). (\( f(x) - f(x^o) > 0 \) for a minimum.)

3. The generalization of the first order condition is as follows:

*Proposition 1.* If a differentiable function \( f \) has a maximum or a minimum at \( x^o \in X \), then \( f_i(x^o) = 0 \), for all \( i \).

The \( n \) equations generated by setting each partial derivative equal to zero represent the first order conditions. If a solution exists, then they may be solved for the \( n \) solution values \( x_i^o \).

4. As in the case of \( n \) choice variables, there are second order conditions which determine whether a critical point is a maximum or a minimum. The complication is that there is no longer one second order derivative which can be checked for negativity or positivity. In fact, there are \( n^2 \) such derivatives, \( f_{ij}(x^o) \), \( i, j = 1, \ldots, n \). The relevant second order condition for a maximum is that

\[
(*) \quad d^2 f(x^o) = \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(x^o) \ dx_i dx_j < 0 \text{ for all } (dx_1, \ldots, dx_n) \neq 0.
\]

This condition is that the quadratic form \( \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(x^o) \ dx_i dx_j \) is *negative definite.*
In (*), the discriminate is the Hessian matrix of $f$ (the objective function). As discussed above, the rather cumbersome (*) condition is equivalent to a fairly simple sign condition. This is as follows:

\[
\begin{bmatrix}
    f_{11} & \cdots & f_{1n} \\
    \vdots & \ddots & \vdots \\
    f_{n1} & \cdots & f_{nn}
\end{bmatrix}
\]

(SOC) (max) $|PM_i|$ of $H = \begin{bmatrix}
    \cdot & \cdots & \cdot \\
    \cdot & \ddots & \cdot \\
    \cdot & \cdots & \cdot
\end{bmatrix}$, evaluated at $x^o$, have signs $(-1)^i$.

The analogous conditions for a minimum are that

\[
\begin{bmatrix}
    f_{11} & \cdots & f_{1n} \\
    \vdots & \ddots & \vdots \\
    f_{n1} & \cdots & f_{nn}
\end{bmatrix}
\]

(SOC) (min) $|PM_i|$ of $H = \begin{bmatrix}
    \cdot & \cdots & \cdot \\
    \cdot & \ddots & \cdot \\
    \cdot & \cdots & \cdot
\end{bmatrix}$, evaluated at $x^o$, have positive signs

If $f$ satisfies the SOC for a maximum globally, then $f$ is strictly concave. If it satisfies the SOC for a minimum globally, then $f$ is strictly convex. For an $n$ variable function, the definition of strict concavity reads the same: $f(\alpha x + (1-\alpha)x') > \alpha f(x) + (1-\alpha)f(x')$, $x \neq x'$, $\alpha \in (0,1)$.

**Proposition 2.** If at a point $x^o$ we have

(i) $f_i(x^o) = 0$, for all $i$, and

(ii) SOC for a maximum (minimum) is satisfied at $x^o$, 


Then \( x^o \) is a local maximum (minimum). If in addition the SOC is met for all \( x \in X \) or if \( f \) is strictly concave (convex), then \( x^o \) is a unique global maximum (minimum).

**Examples:** #1 Maximizing a profit function over two strategy variables. Let profit be a function of the two variables \( x_i \), \( i = 1, 2 \). The profit function is \( \pi(x_1, x_2) = R(x_1, x_2) - \Sigma r_i x_i \), where \( r_i \) is the unit cost of \( x_i \) and \( R \) is revenue. We wish to characterize a profit maximal choice of \( x_i \). The problem is written as

\[
\max_{(x_1, x_2)} \pi(x_1, x_2).
\]

The FOC are

\[
\pi_1(x_1, x_2) = 0 \\
\pi_2(x_1, x_2) = 0.
\]

The second order conditions are

\[
\pi_{11} < 0, \pi_{11}\pi_{22} - \pi_{12}^2 > 0 \text{ (recall Young’s Theorem } \pi_{ij} = \pi_{ji}).
\]

The effect of a change in \( r_1 \) can be determined by differentiating the FOC with respect to \( r_1 \). We obtain

\[
H \left[ \frac{\partial x_1}{\partial r_1} / \frac{\partial r_1}{\partial r_1} \right] = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{ where } H \text{ is the relevant Hessian.}
\]

Using Cramer’s rule,

\[
\frac{\partial x_1}{\partial r_1} = \frac{1 \ \pi_{12}}{|H|} = \frac{\pi_{22}}{|H|} < 0.
\]
Likewise
\[
\frac{\partial x_2}{\partial x_1} = \begin{vmatrix} \pi_{11} & 1 \\ \pi_{21} & 0 \\ \end{vmatrix} = \frac{-\pi_{21}}{|H|}.
\]

The sign of \(\pi_{21}\) is positive if 1 and 2 are complements in profit and it is negative if they are substitutes.

#2. \(\min_{(x,y)} x^2 + xy + 2y^2\). The FOC are

\[
2x + y = 0,
\]

\[
x + 4y = 0.
\]

Solving for the critical values \(x = 0\) and \(y = 0\). \(f_{11} = 2, f_{12} = 1\) and \(f_{22} = 4\). The Hessian is

\[
H = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}, \text{ with } f_{11} = 2 > 0 \text{ and } |H| = 8 - 1 = 7 > 0.
\]

Thus, \((0,0)\) is a minimum. Further, it is global, because the Hessian sign conditions are met for any \(x,y\).

**Existence**

In the case of a function of many variables, we want to generalize our existence argument above. To do this, we must introduce a few concepts.

**Def. 1.** A set \(X \subset \mathbb{R}^N\) is said to be *open* if for all \(x \in X \exists N(x)\) such that \(N(x) \subset X\). The set \(X\) is said to be *closed* if its complement is open.
Def. 2. A set $X \subseteq \mathbb{R}^N$ is said to be bounded if the distance between any two of its points is finite.

That is, $\left[ \sum_{i=1}^{n} (x_i - x'_i)^2 \right]^{1/2} < \infty$, for all $x, x' \in X$.

Def. 3 A set $X \subseteq \mathbb{R}^N$ is said to be compact if it is both closed and bounded.

We can now state a basic existence result.

**Proposition.** Let $f: X \rightarrow \mathbb{R}$, where $X$ is a subset of $\mathbb{R}^N$. If $X$ is compact and $f$ is continuous, then $f$ has a maximum and a minimum on $X$. If $X$ is both compact and convex and $f$ is strictly concave, then $f$ has a unique maximum on $X$. If $X$ is both compact and convex and $f$ is strictly convex, then $f$ has a unique minimum on $X$.

This proposition does not distinguish between boundary optima and interior optima. As in the case of a function of single variable, the results can be used to show the existence of interior optima by showing that boundary optima are dominated. The technique is as described above.

**Constrained Optimization**

1. The basic problem is to maximize a function of at least two independent variables subject to a constraint. We write the objective function as $f(x)$ and the constraint as $g(x) = 0$. The constraint set is written as $C = \{ \ x \ | \ g(x) = 0 \}$. The function $f$ maps a subset of $\mathbb{R}^n$ into the real line. We write the problem as

$$\text{Max} \ f(x) \text{ subject to } g(x) = 0.$$ 

2. A local constrained extremum is defined as follows:

Def. $x^o$ is a local maximum (minimum) of $f(x)$ subject to $g(x) = 0$ if there exists $N(x^o)$ such that $N(x^o) \cap C \neq \emptyset$ and $\forall \ x \in N(x^o) \cap C, f(x) < (>) f(x^o)$. 

3. The basic result is explained in

*Proposition 1.* Let \( f \) be a differentiable function whose \( n \) independent variables are restricted by the differentiable constraint \( g(x) = 0 \). Form the function \( L(\lambda, x) \equiv f(x) + \lambda g(x) \), where \( \lambda \) is an undetermined multiplier. If \( x^o \) is an interior maximizer or minimizer of \( f \) subject to \( g(x) = 0 \), then there is a \( \lambda^o \) such that

(1) \( \frac{\partial L(\lambda^o, x^o)}{\partial x_i} = 0 \), for all \( i \), and

(2) \( \frac{\partial L(\lambda^o, x^o)}{\partial \lambda} = 0 \).

Remark: \( L \) is the Lagrangian function and \( \lambda \) is the Lagrangian multiplier.

4. The relevant SOC for a maximum is that

\[ (*) \quad d^2f(x^o) = \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(x^o) \, dx_i dx_j < 0 \quad \text{for all } (dx_1, \ldots, dx_n) \neq 0 \quad \text{such that } dg = 0. \]

Condition \( (*) \) says that the quadratic form \( d^2f \) is negative definite subject to the constraint that \( dg = 0 \). This is equivalent to \( d^2L = d^2f \) being negative definite subject to \( dg = 0 \), because \( dL = df + \lambda dg + gd\lambda = df \), if \( dg = g = 0 \). Condition \( (*) \) is equivalent to a rather convenient condition involving the bordered Hessian matrix. The bordered Hessian is given by

\[
\overline{H}(\lambda^o, x^o) = \begin{bmatrix}
0 & g_1 & \cdots & g_n \\
g_1 & L_{11} & \cdots & L_{1n} \\
& \ddots & \ddots & \ddots \\
& & \ddots & \ddots \\
g_n & L_{n1} & \cdots & L_{nn}
\end{bmatrix}
\]

The sign condition is
SOC (max) \[ |PM_i| \text{ of } |\overline{H}| \text{ of order } i \geq 3, \text{ evaluated at } (\lambda^o, x^o), \text{ has sign } (-1)^{i+1}. \]

For a minimum, the second order condition is

\[ (** \right) \quad d^2 f(x^o) = \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(x^o) \, dx_i \, dx_j > 0 \text{ for all } (dx_1, ..., dx_n) \neq 0 \text{ such that } dg = 0. \]

and the equivalent sign condition is

SOC (min) \[ |PM_i| \text{ of } |\overline{H}| \text{ of order } i \geq 3, \text{ evaluated at } (\lambda^o, x^o), \text{ are all negative.} \]

5. The sufficiency result is

Proposition 2. Let \( f \) be a differentiable function whose \( n \) independent variables are restricted by the differentiable constraint \( g(x) = 0 \). Form the function \( L(\lambda, x) \equiv f(x) + \lambda g(x) \), where \( \lambda \) is an undetermined multiplier. Let there exist an \( x^o \) and a \( \lambda^o \) such that

(1) \[ \partial L(\lambda^o, x^o)/\partial x_i = 0, \text{ for all } i, \text{ and} \]

(2) \[ \partial L(\lambda^o, x^o)/\partial \lambda = 0. \]

Then \( x^o \) is a local maximum (minimum) of \( f(x) \) subject to \( g(x) = 0 \), if, in addition to (1) and (2) the SOC for a maximum (minimum) is satisfied. If SOC is met for all \( x \in C \), then \( x^o \) is a unique global maximizer (minimizer).

In the above maximization problem, as long as the relevant constraint set is convex, the maximum will be a global maximum if the objective function is \textit{strictly quasi-concave}. The latter property means that the sets \{ \( x : f(x) \geq f(x') \) \} = \( U^i(x') \) are strictly convex.\(^1\) The latter sets are called upper level sets. If the lower level sets of a function are strictly convex sets, then the

\(^1\) A set \( X \) in \( \mathbb{R}^n \) is strictly convex if \( \alpha x + (1-\alpha)x' \in \text{interior } X \) for all \( \alpha \in (0,1) \) and all \( x \neq x' \) with \( x, x' \in X \).
function is said to be strictly quasi-convex. When the relevant constraint set is convex and the objective function is strictly quasi-convex (strict quasi-concave), then the minimum (maximum) is unique. A bordered Hessian condition may be used to check for either of these properties. We consider this next.

If a function is defined on the nonnegative orthant of \( \mathbb{R}^N \) and is twice continuously differentiable, then there is an easy operational check for quasi-concavity or quasi-convexity. Let \( f(x_1, \ldots, x_N) \) be such a function. The function \( f \) is strictly quasi-concave if

\[
B = \begin{bmatrix}
0 & f_1 & f_2 & \cdots & f_N \\
 f_1 & f_{11} & f_{12} & \cdots & f_{1N} \\
f_2 & f_{12} & f_{22} & \cdots & f_{2N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
f_N & f_{1N} & f_{2N} & \cdots & f_{NN}
\end{bmatrix}
\]

has \( |PM_i| \) with sign \( (-1)^{i+1} \) for \( i \geq 2 \).

For strict quasi-convexity, the condition is that \( |PM_i| \) are all negative for \( i \geq 2 \). It is important to note that the condition pertains only to functions defined on \( \mathbb{R}_+^N = \{ x \in \mathbb{R}^N : x_i \geq 0 \text{ for all } i \} \). The condition on \( B \) is equivalent to the statement that \( d^2f \) is negative definite subject to \( df = 0 \).

**Existence**

For the present problem, the conditions for existence of a constrained optimum are the same as those for the unconstrained problem except that the set \( X \cap C \) is the relevant set for which we assume boundedness and closedness (compactness). Further, the objective function is assumed to be continuous over this set. Under these conditions, a constrained optimum will exist.
Extension to Multiple Constraints

Suppose that there are \( m < n \) constraints \( g_j(x) = 0, \ j = 1, \ldots, m \). The Lagrangian is written as

\[ L(\lambda_1, \ldots, \lambda_m, x_1, \ldots, x_n) = f(x) + \sum \lambda_j g_j(x). \]

The FOC are that the derivatives of \( L \) in \( x_i \) and \( \lambda_j \) vanish:

\[ f_i + \sum \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \text{ for all } i, \]

\[ g_j(x) = 0, \text{ for all } j. \]

The bordered Hessian becomes

\[
|\tilde{H}| = \begin{bmatrix}
0 & J_g \\
J_g & L_g
\end{bmatrix},
\]

where \( J_g \) is the Jacobian of the constraint system in \( x \), and \( [L_{ij}] \) is the Hessian of the function \( L \) in \( x \). The condition \( m < n \) must be met, and the sign conditions for a maximum and a minimum are written in terms of the principal minors of the above bordered Hessian. For a maximum, the condition is

(SOC)(max) \ PM_i \ of \ |\tilde{H}| \ of \ order \ i > 2m \ has \ sign \ (-1)^r, \ where \ r \ is \ the \ order \ of \ the \ largest \ order square \ [L_{ij}] \ embedded \ in \ |PM_i|.

For a minimum, the condition is

(SOC)(min) \ |PM_i| \ of \ |\tilde{H}| \ of \ order \ i > 2m \ has \ sign \ (-1)^m.

Examples:

#1. Find the critical 4-tuple for the function \( y = f(x_1, x_2, x_3), \ f: \mathbb{R}_+^3 \to \mathbb{R} \)

\[ y = x_1 x_2 x_3 \]
subject to

\[ x_1^2 + x_2^2 - 1 = 0 \]

\[ x_1 - x_3 = 0. \]

Restrict your choice of \( x_i, \ i = 1, 2, 3 \), to the positive reals.

\[ L = x_1 x_2 x_3 + \lambda_1 (x_1^2 + x_2^2 - 1) + \lambda_2 (x_1 - x_3) \]

(1) \[ x_2 x_3 + 2\lambda_1 x_1 + \lambda_2 = 0 \]

(2) \[ x_1 x_3 + 2\lambda_1 x_2 = 0 \]

(3) \[ x_1 x_2 - \lambda_2 = 0 \]

(4) \[ x_1^2 + x_2^2 - 1 = 0 \]

(5) \[ x_1 - x_3 = 0 \]

From (5)

\[ x_1 = x_3 \]

From (3)

\[ x_1 x_2 = x_3 x_2 = \lambda_2 \]

Solve for \( \lambda_1 \) from 2

\[ x_1 x_3 + 2\lambda_1 x_2 = 0 \]

\[ x_1^2 = -2\lambda_1 x_2 \]

\[ \therefore \lambda_1 = \frac{-x_1^2}{2x_2} \]

Go to (1)

\[ x_1 x_2 + 2 \left( \frac{-x_1^2}{2x_2} \right) x_1 + x_1 x_2 = 0 \]

\[ 2x_1 x_2 - \frac{x_1^3}{x_2} = 0 \]
Multiply by $x_2/x_1$ (both sides)

\[(\alpha) \quad 2x_2^2 - x_1^2 = 0\]

Use (4)

\[(\beta) \quad x_1^2 + x_2^2 - 1 = 0\]

Solve $x_1, x_2$:

\[(\alpha) \quad x_1^2 = 2x_2^2\]

\[\therefore \quad 2x_2^2 + x_2^2 = 1\]

\[3x_2^2 = 1\]

\[x_2^2 = 1/3\]

\[x_2 = \sqrt[3]{3}\]

but

\[x_1^2 = 2x_2^2 = 2(1/3) = 2/3\]

\[x_1 = \sqrt{\frac{2}{3}}\]

\[x_3 = \sqrt{\frac{2}{3}}\]

and

\[y^0 = (1/3)^{1/2}(2/3)^{1/2}(2/3)^{1/2}\]

\[y^0 = \frac{2}{3\sqrt{3}}\]

\[\therefore \quad (\sqrt{2/3}, \sqrt{2/3}, \sqrt{2/3}, \sqrt{3}).\]

The Problem’s Bordered Hessian.

\[
\begin{bmatrix}
0 & 0 & 2x_1 & 2x_2 & 0 \\
0 & 0 & 1 & 0 & -1 \\
2x_1 & 1 & 2\lambda_1 & x_3 & x_2 \\
2x_2 & 0 & x_3 & 2\lambda_1 & x_1 \\
0 & -1 & x_2 & x_1 & 0
\end{bmatrix}
\]
\( x^0 \) represents a max if \( |\bar{H}| = |\bar{H}| \) has sign \((-1)^3 < 0\).

\( x^0 \) represents a min if \( |\bar{H}| = |\bar{H}| \) has sign \((-1)^2 > 0\).

\#2. Let \( y = x_1^{1/3} x_2^{1/3}, f: \mathbb{R}^2 \to \mathbb{R} \)

\[ 3x_1 + 3x_2 - 30 = 0 \]

\[ L = x_1^{1/3} x_2^{1/3} + \lambda (3x_1 + 3x_2 - 30) \]

(1) \[ \frac{1}{3y} \frac{x_1}{x_1} + 3\lambda = 0 \]

(2) \[ \frac{1}{3y} \frac{x_2}{x_2} + 3\lambda = 0 \]

(3) \[ 3x_1 + 3x_2 - 30 = 0 \]

Solve: Divide (1) by (2)

\[ \frac{1/3y}{x_1} = \frac{1/3y}{x_2} \frac{x_2}{x_1} = \frac{-3\lambda}{3\lambda} = +1 \]

\[ \frac{x_2}{x_1} = +1 \]

\[ x_2 = +x_1 \]

From (3)

\[ 3x_1 + 3x_1 = 30 \]

\[ 6x_1 = 30 \]

\[ x_1^0 = 5 \]

\[ x_2^0 = 5 \]

\[ y^0 = (5)^{1/3} (5)^{1/3} \]

\[ y^0 = 5^{2/3} \]

Critical triple is
\[(5, \ 5, \ 5^{2/3})\]

Find \( [\mathbb{H}] \)

\[
[\mathbb{H}] = \begin{bmatrix}
0 & 3 & 3 \\
3 & f_{11}^0 & f_{12}^0 \\
3 & f_{21}^0 & f_{22}^0
\end{bmatrix}
\]

\[
f_1 = \frac{1/3y}{x_1} = \frac{1/3x_1^{1/3}x_2^{1/3}}{x_1}
\]

\[
= \frac{1}{3}x_1^{-2/3}x_2^{1/3}
\]

\[
f_{11} = \frac{1}{3}\left(\frac{-2}{3}\right)x_1^{-5/3}x_2^{1/3}
\]

\[
= \frac{-2}{9}x_1^{-5/3}x_2^{1/3}
\]

\[
f_{12} = f_{21} = \frac{1}{3}x_1^{-2/3}\left(\frac{1}{3}\right)x_2^{-2/3}
\]

\[
= \frac{1}{9}x_1^{-2/3}x_2^{-2/3}
\]

\[
f_{22} = \frac{1}{3}x_1^{1/3}x_2^{-2/3}
\]

\[
f_{22} = \frac{-2}{9}x_1^{1/3}x_2^{-5/3}
\]

Hence;

\[
[\mathbb{H}] = \begin{bmatrix}
0 & 3 & 3 \\
3 & -\frac{2}{9}x_1^{-5/3}x_2^{1/3} & \frac{1}{9}x_1^{-2/3}x_2^{-2/3} \\
3 & \frac{1}{9}x_1^{2/3}x_2^{-2/3} & -\frac{2}{9}x_1^{1/3}x_2^{-5/3}
\end{bmatrix}
\]

For a max \([\mathbb{H}]_0\) has sign \((-1)^3 > 0.\)

Expand \([\mathbb{H}]:\)
\[ [H] = -3 \left[ \left(-\frac{2}{3}x_1^{1/3}x_2^{-5/3}\right) - 3 \left[ \frac{2}{3}x_1^{-2/3}x_2^{-2/3}\right] \right] + 3 \left[ 3 \left( \frac{2}{3}x_1^{-2/3}x_2^{-2/3}\right) -3\left(-\frac{2}{3}x_1^{-5/3}x_2^{1/3}\right) \right] \]
\[ = 9 \frac{2}{3}x_1^{1/3}x_2^{-5/3} + 9 \frac{2}{3}x_1^{-2/3}x_2^{-2/3} + 9 \frac{2}{3}x_1^{-2/3}x_2^{-2/3} + 9 \frac{2}{3}x_1^{-5/3}x_2^{1/3} \]
\[ = 2 x_1^{1/3}x_2^{-5/3} + 2 x_1^{-2/3}x_2^{-2/3} + 2 x_1^{-5/3}x_2^{1/3} \]
\[ [H] = 2 \left[ x_1^{1/3}x_2^{-5/3} + x_1^{-2/3}x_2^{-2/3} + x_1^{-5/3}x_2^{1/3} \right] \]

Now must evaluate 2nd partials at \( x^0 \):
\[ [H] = 2\left[ 5^{1/3}5^{-5/3} + 5^{-2/3}5^{-2/3} + 5^{-5/3}5^{1/3} \right] \]
\[ = 2 \left[ 5^{-4/3} + 5^{-4/3} + 5^{-4/3} \right] \]
\[ = 2 \left[ 3 \cdot 5^{-4/3} \right] \]
\[ = 6 \left( 5 \right)^{-4/3} = \frac{6}{5^{4/3}} > 0 \]

Hence a max at
\( (5, 5, 5^{2/3}) \).

**Inequality Constrained Problems**

1. In many problems the side constraints are is best represented as inequality constraints:

\[ g_j(x) \geq 0. \]

We wish to characterize the FOC to this extended problem.

\[ \text{Max} \ f(x) \text{ subject to } g_j(x) \geq 0, \ j = 1, \ldots, m. \]

2. Form the Lagrangian

\[ L(\lambda, x) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x). \]
The FOC or necessary conditions are as follows:

(1) \[ \frac{\partial L}{\partial x_i} = f_i + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x)}{\partial x_i} = 0 \text{ for all } i, \]

(2) \[ \frac{\partial L}{\partial \lambda_j} = g_j \geq 0, \lambda_j \geq 0 \text{ and } g_j \lambda_j = 0, j = 1, \ldots, M \text{ and} \]

(3) The Constraint Qualification holds.

Remark 1: These same conditions are used for a minimum. If one were to solve Min f(x) subject to \( g_j(x) \leq 0 \), then to use the above conditions one would rewrite the problem as Max -f(x) subject to \(-g_j(x) \geq 0\).

Remark 2: The FOC (1) and (2) are necessary conditions only if the Constraint Qualification holds. This rules out particular irregularities by imposing restrictions on the boundary of the feasible set. These irregularities would invalidate the FOC (1) and (2) should the solution occur there. Let \( x^o \) be the point at which (1) and (2) hold and let index set \( k = 1, \ldots, K \) represent the set of \( g_j \) which are satisfied with equality at \( x^o \). Then the matrix

\[
J = \begin{bmatrix}
\frac{\partial g_1(x^o)}{\partial x_1} & \ldots & \frac{\partial g_1(x^o)}{\partial x_n} \\
\frac{\partial g_2(x^o)}{\partial x_1} & \ldots & \frac{\partial g_2(x^o)}{\partial x_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial g_K(x^o)}{\partial x_1} & \ldots & \frac{\partial g_K(x^o)}{\partial x_n}
\end{bmatrix}
\]

Has rank \( K \leq n \). That is the gradient vectors of the set of equality constraints are linearly independent.
3. Example: Nonnegativity constraints on \( x_i \)

\[
L = f + \lambda_1 x_1 + \lambda_2 x_2.
\]

The FOC are

\[
f_i + \lambda_i = 0
\]

\[
x_i \geq 0, \lambda_i \geq 0 \text{ and } \lambda_i x_i = 0.
\]

If all \( x_i > 0 \), then \( \lambda_i = 0 \) and we have the previous conditions.

If \( \lambda_i > 0 \), then it must be true that \( x_i = 0 \) and \( f_i < 0 \) at the optimum.

It can be true that \( \lambda_i = 0 \) and \( x_i = 0 \).

4. Example: Consider the problem where we seek to Max \( f(x) \), \( x \in \mathbb{R}^2 \), subject to \( p_1 x_1 + p_2 x_2 \leq c \), where \( c \) is a constant and \( x_i \geq 0 \). Assume that \( c, p_i > 0 \), We should set up the Lagrangian as

\[
L = f(x) + \lambda [c - p_1 x_1 - p_2 x_2] + \Sigma \gamma_i x_i.
\]

The relevant first order conditions are as follows:

\[
f_i - \lambda p_i + \gamma_i = 0, \ i = 1,2,
\]

\[
c - p_1 x_1 - p_2 x_2 \geq 0, \lambda \geq 0, \lambda [c - p_1 x_1 - p_2 x_2] = 0, \text{ and}
\]

\[
x_i \geq 0, \gamma_i \geq 0, \gamma_i x_i = 0, i = 1,2.
\]

Note that if the constraints are nonbinding and, at optimum, it is true that \( [c - p_1 x_1 - p_2 x_2], x_i > 0 \), then the optimum is defined as a free optimum with \( f_i = 0 \). Next, suppose that the constraint \( p_1 x_1 \)
+ p_2x_2 \leq c$ is binding, $\gamma_2 > 0$, and that $x_1 > 0$. In this case, $x_2 = 0$, $\gamma_1 = 0$, and $p_1x_1 + p_2x_2 - c = 0$ and the first order conditions read

$$f_1 = \lambda p_1$$

$$f_2 = \lambda p_2 - \gamma_2$$

which implies that $f_2 < \lambda p_2$.

Thus, we have that

$$\frac{f_1}{f_2} > \frac{p_1}{p_2}, \ c = p_1x_1 \Rightarrow x_1 = c/p_1 \ \text{and} \ x_2 = 0.$$

The illustration is as follows:

If it were true that $\gamma_i = 0$ and $\lambda > 0$, then

$$\frac{f_1}{f_2} = \frac{p_1}{p_2}, \ c = p_1x_1 + p_2x_2, \ \text{and} \ x_1, x_2 > 0.$$

This is an interior optimum. The illustration is as follows:
\[ \text{slope} = \frac{p_1}{p_2} = \frac{f_1}{f_2} \]