Mathematics Tools

Advanced Microeconomics, Fall 2016

Menghan Xu

WISE & SOE
Xiamen University

Mathematical tools

"Mathematics is language" by Paul Samuelson.

- More precise description of assumptions
  eg., utility theory
- Rigorous logic in reasoning
  eg., everyday examples against rules of basic logics
- Giving results that are not obvious
  eg., First Welfare Theorem, Coase Theorem

Basic logic: conditional statement

- Start from "If event **A** then event **B**."
    - **A** is a sufficient condition for **B**
    - **B** is a necessary condition for **A**
    - **A** only if **B**
    - **A** implies **B**
    - **B** is implied by **A**
    - **A** $\Rightarrow$ **B**
    - **B** $\Leftarrow$ **A**
    - If **A** $\Rightarrow$ **B**, then (not **B**) $\Rightarrow$ (not **A**)

Basic logic: conditional statement

- Start from "If event **A** then event **B**."
  - **A** is a sufficient condition for **B**
  - **B** is a necessary condition for **A**
  - **A** only if **B**
  - **A** implies **B**
  - **B** is implied by **A**
  - **A** $\Rightarrow$ **B**
  - **B** $\Leftarrow$ **A**
  - If **A** $\Rightarrow$ **B**, then (not **B**) $\Rightarrow$ (not **A**)
  - (not **A**) $\Rightarrow$? **B**
    - "if you take microeconomics then you will become wiser"
      It is unlikely to mean "if you don't take microeconomics then you won't become wiser"

Basic logic: conditional statement

- Start from "If event **A** then event **B**."
  - **A** is a sufficient condition for **B**
  - **B** is a necessary condition for **A**
  - **A** only if **B**
  - **A** implies **B**
  - **B** is implied by **A**
  - $\mathbf{A} \Rightarrow \mathbf{B}$
  - $\mathbf{B} \Leftarrow \mathbf{A}$
  - If $\mathbf{A} \Rightarrow \mathbf{B}$, then (not $\mathbf{B}$) $\Rightarrow$ (not $\mathbf{A}$)
  - (not $\mathbf{A}$) $\Rightarrow$? $\mathbf{B}$
    - "if you take microeconomics then you will become wiser"
      It is unlikely to mean "if you don't take microeconomics then you won't become wiser"
    - "if it is sunny tomorrow then let's go hiking"
      It probably means "if it is not sunny tomorrow then let's not go hiking"

Basic logic: conditional statement

- Start from "If event **A** then event **B**."
  - **A** is a sufficient condition for **B**
  - **B** is a necessary condition for **A**
  - **A** only if **B**
  - **A** implies **B**
  - **B** is implied by **A**
  - $\mathbf{A} \Rightarrow \mathbf{B}$
  - $\mathbf{B} \Leftarrow \mathbf{A}$
  - If $\mathbf{A} \Rightarrow \mathbf{B}$, then (not **B**) $\Rightarrow$ (not **A**)
  - (not **A**) $\Rightarrow$? **B**
    - "if you take microeconomics then you will become wiser"
      It is unlikely to mean "if you don't take microeconomics then you won't become wiser"
    - "if it is sunny tomorrow then let's go hiking"
      It probably means "if it is not sunny tomorrow then let's not go hiking"
  - Statement "not (**A** and **B**)" is equivalent to "(not **A**) or (not **B**)"

Interval and set of vectors

- Interval: a set of (real) numbers between (and probably including) two numbers
- Infinity ∞: for intervals that extend indefinitely in one or both directions
- Interior of an interval: the set of all numbers in the interval except the endpoints

Interval and set of vectors

- Interval: a set of (real) numbers between (and probably including) two numbers
- Infinity ∞: for intervals that extend indefinitely in one or both directions
- Interior of an interval: the set of all numbers in the interval except the endpoints
- A point $\mathbf{x}$ is a **boundary point** of a set S of vectors if for *every* number $\epsilon > 0$, at least one point within the distance ϵ of $\mathbf{x}$ is in S, and at least one point within the distance ϵ of $\mathbf{x}$ is outside S.
- A point $\mathbf{x}$ is an **interior point** of a set S of vectors if there is a number $\epsilon > 0$, such that *all* points within the distance ϵ of $\mathbf{x}$ are members of S.
- The set S is **open** if every point in S is an interior point of S.
- The set S is **closed** if every boundary point of S is a member of S.

Function

- Function: a **rule** that associates every point in some set (**domain**) with a point in another set (**range**).
    - Function of a single variable: domain is a set of 1-vectors
    - Function of many variables: domain is a set of n-vectors
    - Real-valued function: when the second set is of real numbers
    - $f : A \rightarrow B$: rule is f, domain is set A, range is a subset of B

Continuous function

- Let f be a function of many variables and let $\mathbf{a}$ be a point in its domain. Then f is continuous at $\mathbf{a}$ if, for any number $\epsilon > 0$, there is a number $\delta > 0$ such that for any value of $\mathbf{x}$ in the domain of f for which $d(\mathbf{x}, \mathbf{a}) < \delta$,$d(f(\mathbf{x}), f(\mathbf{a})) < \epsilon$.
- A function is called a continuous function if it is continuous at every point in its domain.

Properties of continuous function

- If functions f and g of many variables are continuous at $\mathbf{x}_0$, then all the following functions are continuous at x_0
  - $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$
  - $h(\mathbf{x}) = f(\mathbf{x}) * g(\mathbf{x})$
  - $h(\mathbf{x}) = f(\mathbf{x})/g(\mathbf{x})$ for all $\mathbf{x}$ with $g(\mathbf{x}) \neq 0$

Properties of continuous function

- If functions f and g of many variables are continuous at $\mathbf{x}_0$, then all the following functions are continuous at x_0
  - $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$
  - $h(\mathbf{x}) = f(\mathbf{x}) * g(\mathbf{x})$
  - $h(\mathbf{x}) = f(\mathbf{x})/g(\mathbf{x})$ for all $\mathbf{x}$ with $g(\mathbf{x}) \neq 0$
- All polynomials are continuous
  Polynomial: $f(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + ... + \alpha_k x^k$ where k is a nonnegative integer and αs are numbers

More on continuous function

- Composition of continuous functions: if the function f of many variables is continuous at $\mathbf{x}_0$ and the function g of a single variable is continuous at $f(\mathbf{x}_0)$, then function $h(\mathbf{x}) = g(f(\mathbf{x}))$ for all $\mathbf{x}$ is continuous at $\mathbf{x}_0$

More on continuous function

- Composition of continuous functions: if the function f of many variables is continuous at $\mathbf{x}_0$ and the function g of a single variable is continuous at $f(\mathbf{x}_0)$, then function $h(\mathbf{x}) = g(f(\mathbf{x}))$ for all $\mathbf{x}$ is continuous at $\mathbf{x}_0$
- Intermediate value theorem: if f is a continuous function of a single variable with domain $[a, b]$ and $a \neq b$, then for any number y between $f(a)$ and $f(b)$, or equal to $f(a)$ or $f(b)$, $\exists x \in [a, b]$ such that $f(x) = y$.

Univariate function calculus: differentiation

- Differentiable: the function f of a single variable defined on an open interval is **differentiable** at a if $lim_{h\to 0}\frac{f(a+h)-f(a)}{h}$ exists, in which case this limit is the **derivative of the function f at a**, denoted $f'(a)$.
- Rules of differentiation
- The derivative of a function of a single variable at a point is a good linear approximation of the function around the point. If no good linear approximation exists at some point x (eg., when the graph has a "kink" at point x), the function is not differentiable at x.

More on differentiation

- Chain rule: if f and g are differentiable functions of a single variable and the function F is defined by $F(x) = f(g(x))$, then

$$F'(x) = f'(g(x))g'(x)$$

- If f' is differentiable at a then f is twice-differentiable, and the derivative at a is called the second derivative of f at a, denoted by f''

Integration

Definite integral: let f be a function of a single variable on the domain $[a, b]$, the definite integral of f from a to b, denoted by

$$\int_a^b f(z)dz$$

is a measure of the area between the horizontal axis and the graph of f, between a and b.

Fundamental theorem of calculus

Let f be an integrable function of a single variable defined on $[a, b]$. Define the function F of a single variable on the domain $[a, b]$ by

$$F(x) = \int_a^x f(z)dz$$

If f is continuous at the point c in $[a, b]$, then F is differentiable at c and

$$F'(c) = f(c)$$

Similarly, define the function G on $[a, b]$ by $G(x) = \int_x^b f(z)d_z$. If f is continuous at the point c in $[a, b]$, the G is differentiable at c and

$$G'(c) = -f(c)$$

If f is continuous on $[a, b]$ and $f = F'$ for some function F, then

$$\int_a^b f(z)d_z = F(b) - F(a)$$

Multivariate function calculus: partial derivative

- Partial derivative: function $f : D \to \Re$ where D is a subset of $\Re^n$, then the partial derivative of f with respect to x_i at x is

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, ..., x_i + h, ..., x_n) - f(x_1, ..., x_i, ..., x_n)}{h}$$

  which can also be denoted as $f_i(\mathbf{x})$
- If all the partial derivatives of f exist and are continuous functions, then f is **continuously differentiable**.
- Gradient vector of f at $\mathbf{x}$: a column vector of all n partial derivatives $\nabla f(\mathbf{x}) = (f_1(\mathbf{x}), ..., f_n(\mathbf{x}))'$. ∇ is pronounced as "*del*".
- Directional derivative: $\nabla f(\mathbf{x})'\mathbf{z}$, or $\sum_{i=1}^{n} f_i(\mathbf{x}) z_i$, is the rate of change of f moving at a rate and direction given by $\mathbf{z} = (z_1, ..., z_n)'$ at point $\mathbf{x}$.

$$\nabla f(\mathbf{x})'\mathbf{z} = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{z}) - f(\mathbf{x})}{t}$$

Second order derivative

- Second order partial derivative: $f_{ij}(\mathbf{x}) = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}(\frac{\partial f(\mathbf{x})}{\partial x_i})$
  Cross partials: the derivative of f_i with respect to its jth argument evaluated at $(x_1, ..., x_n)'$ is called a "cross partial".
- Hessian matrix $H_f(\mathbf{x})$: taking gradient of the gradient and arrange them into an n by n matrix with $f_{ij}(\mathbf{x})$ the (i, j)th element.
- Young's theorem: Let f be a differentiable function of n variables. If each of the cross partials f_{ij} and f_{ji} exists and is continuous at all points in some open set S of values of $(x_1, ..., x_n)'$ then $f_{ij}(x_1, ..., x_n) = f_{ji}(x_1, ..., x_n)$ for all $(x_1, ..., x_n)'$ in S.

Taylor approximation

- A function of many variables is differentiable at a point if there exists a good linear approximation of the function around the point.
- Taylor approximation: if function $f : \Re^n \to \Re$ is twice differentiable then

$$f(\mathbf{x}+\mathbf{h}) \approx f(\mathbf{x}) + \nabla f(\mathbf{x})'\mathbf{h} + \frac{1}{2!}\mathbf{h}'H(\mathbf{x})\mathbf{h}$$

Chain rule

Chain rule (a simple case): suppose g and h are differentiable functions of a single variable, f is a differentiable function of two variables, and the function F of a single variable is defined by

$$F(x) = f(g(x), h(x))$$

then

$$F'(x) = f_1(g(x), h(x))g'(x) + f_2(g(x), h(x))h'(x)$$

Chain rule, continue

Chain rule: if g^j is a differentiable function of m variables for $j = 1, 2, ...n$, f is a differentiable function of n variables, and the function F of m variables is defined by

$$F(x_1, ..., x_m) = f(g^1(x_1, ..., x_m), ..., g^n(x_1, ..., x_m)) \text{ for all } (x_1, ..., x_m)$$

then

$$F_j(x_1, ..., x_m) = \sum_{i=1}^{n} f_i(g^1(x_1, ..., x_m), ..., g^n(x_1, ..., x_m)){g^i}_j(x_1, ..., x_m)$$

where ${g^i}_j$ is the partial derivative of g^i with respect to its jth argument.

Integration

Leibniz's formula: let f be a differentiable function of two variables, let a and b be differentiable functions of a single variable, and define the function F by

$$F(t) = \int_{a(t)}^{b(t)} f(t,x)d_x \text{ for all } t$$

then

$$F'(t) = f(t,b(t))b'(t) - f(t,a(t))a'(t) + \int_{a(t)}^{b(t)} f_1(t,x)d_x$$

Use chain rule and the fundamental theorem of calculus

Implicit function

- $f(x, p) = 0$, where p is a parameter and x is a variable we are interested in, defines x implicitly as a function of p.
- Sometimes we can solve explicitly for x as a function of p, sometimes we can not. But we still want to know how equilibrium value of x depends on parameter p (**comparative static analysis**, assuming this equation of x has a solution).
- Rewrite the implicit function as $f(x(p), p) = 0$
- Taking derivative with respect to p

$$f_1(x(p), p)x'(p) + f_2(x(p), p) = 0$$
$$x'(p) = -\frac{f_2(x(p), p)}{f_1(x(p), p)}$$

Homogeneous function

- Homogeneous of degree k: a function f of n variables for which $(tx_1, ..., tx_n)$ is in the domain whenever $t > 0$ and $(x_1, ..., x_n)$ is in the domain is homogeneous of degree k if

  $f(tx_1, ...tx_n) = t^k f(x_1, ..., x_n)$ in the domain of f and all $t > 0$

Homogeneous function

- Homogeneous of degree k: a function f of n variables for which $(tx_1, ..., tx_n)$ is in the domain whenever $t > 0$ and $(x_1, ..., x_n)$ is in the domain is homogeneous of degree k if

  $f(tx_1, ...tx_n) = t^k f(x_1, ..., x_n)$ in the domain of f and all $t > 0$

- Let f be a differentiable function of n variables that is homogeneous of degree k. Then each of its partial derivatives f_i (for i=1,...,n) is homogeneous of degree $k - 1$.

  Example: homogeneous utility/production function, expansion path

Euler's theorem

- The differentiable function f of n variables is homogeneous of degree k if and only if

$$\sum_{i=1}^{n} f_i(x_1, ..., x_n)x_i = kf(x_1, ..., x_n) \text{ for all } (x_1, ...x_n)$$

or in vector form $\nabla f(\mathbf{x})'\mathbf{x} = kf(\mathbf{x})$ for all $\mathbf{x}$

- *Example:*
    - *production function homogenenous of degree one*

$$MP_k * K + MP_L * L = Q$$
$$\frac{r}{MP_K} = \frac{w}{MP_l} = MC$$
$$MC(Q) = AC(Q)$$

    - *production function homogeneous of degree k: average cost vs. marginal cost*

$$MC(Q) = \frac{1}{k}AC(Q)$$

Homothetic function

- For homogeneous utility function $U(x_1, x_2)$, we have $MRS^U(x_1, x_2) = MRS^U(tx_1, tx_2)$.

Homothetic function

- For homogeneous utility function $U(x_1, x_2)$, we have $MRS^U(x_1, x_2) = MRS^U(tx_1, tx_2)$.
- Define function $V(x_1, x_2)$ to be a monotone transformation of the original function $U(x_1, x_2)$, i.e, $V(x_1, x_2) = g(U(x_1, x_2))$ (or $V = g \circ U$) where $g : \Re \to \Re$ is a strictly increasing function.
    - $V(x_1, x_2)$ is not necessarily homogeneous - **cardinal** property
    - $MRS^V(x_1, x_2) = MRS^V(tx_1, tx_2)$ holds - **ordinal** property
- Preference is ordinal so we only want the ordinal property in utility functions, in other words, we only need the *MRS* function to be homogeneous of degree zero!
- Homothetic function: monotone transformation of a homogeneous function (of degree one)
- For production functions, we want to keep the cardinal feature

Quadratic form

- A **quadratic form** in n variables is a function

$$Q(x_1, ...x_n) = b_{11}x_1^2 + b_{12}x_1x_2 + ... + b_{ij}x_ix_j + ... + b_{nn}x_n^2$$

  where b_{ij} for $i = 1, ..., n$ and $j = 1, ..., n$ are constants.
- Matrix equivalence: $Q(\mathbf{x}) = \mathbf{x}'A\mathbf{x}$
  where matrix A is an n by n symmetric matrix with the (i, j)th element $\frac{b_{ij}+b_{ji}}{2}$.

Definiteness

- Let $Q(\mathbf{x})$ be a quadratic form and $Q(\mathbf{x}) = \mathbf{x}'A\mathbf{x}$. A and $Q(\mathbf{x})$ are
  - positive definite if $\mathbf{x}'A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$
  - negative definite if $\mathbf{x}'A\mathbf{x} < 0$ for all $\mathbf{x} \neq 0$
  - positive semidefinite if $\mathbf{x}'A\mathbf{x} \geq 0$ for all $\mathbf{x} \neq 0$
  - negative semidefinite if $\mathbf{x}'A\mathbf{x} \leq 0$ for all $\mathbf{x} \neq 0$
  - indefinite if it is neither positive semidefinite nor negative semidefinite

Necessary and sufficient conditions for definiteness

- A kth order submatrix of matrix A is a **principal** submatrix if it is obtained by deleting $n-k$ rows and the $n-k$ columns with the same numbers.
  Minor is the determinant of a submatrix.
  The kth order **leading principal minor** of A is the kth order minor obtained by deleting the last $n-k$ rows and columns.

Necessary and sufficient conditions for definiteness

- A kth order submatrix of matrix A is a **principal** submatrix if it is obtained by deleting $n-k$ rows and the $n-k$ columns with the same numbers.
  Minor is the determinant of a submatrix.
  The kth order **leading principal minor** of A is the kth order minor obtained by deleting the last $n-k$ rows and columns.
- Let A be an $n \times n$ symmetric matrix and let D_k for $k = 1, ..., n$ be its leading principal minors. Then
  - A is **positive definite** if and only if $D_k > 0$ for $k = 1, ..., n$
  - A is **negative definite** if and only if $(-1)^k D_k > 0$ for $k = 1, ..., n$

Necessary and sufficient conditions for definiteness

- A kth order submatrix of matrix A is a **principal** submatrix if it is obtained by deleting $n-k$ rows and the $n-k$ columns with the same numbers.
  Minor is the determinant of a submatrix.
  The kth order **leading principal minor** of A is the kth order minor obtained by deleting the last $n-k$ rows and columns.
- Let A be an $n \times n$ symmetric matrix and let D_k for $k = 1, ..., n$ be its leading principal minors. Then
  - A is **positive definite** if and only if $D_k > 0$ for $k = 1, ..., n$
  - A is **negative definite** if and only if $(-1)^k D_k > 0$ for $k = 1, ..., n$
- Let A be an $n \times n$ symmetric matrix. Then
  - A is **positive semidefinite** if and only if all the principal minors of A are nonnegative
  - A is **negative semidefinite** if and only if all the kth order principal minors of A are ≤ 0 if k is odd and ≥ 0 if k is even.

Determinants

- $A\mathbf{x}$ is a linear transformation of $\mathbf{x}$,where A is an $n \times n$ matrix and $\mathbf{x}$ is an n-vector. Call the result of the transformation $\mathbf{x}^A$
- Geometric interpretation of the transformation
  - Change of "the area": magnification factor
    - Distance to 0 for 1-dimension
    - Area of a parallelogram for 2-dimension
    - Volume of a parallelepiped for 3-dimension
    - and so on...
  - **Change of the "direction" of the vector $\Rightarrow$ determines the sign of $\mathbf{x}'\mathbf{x}^A$**
  - Determinant is a number that summarizes these two changes

Concave and convex functions

- A set S of n vectors is convex if

$$(1-\lambda)\mathbf{x}+\lambda\tilde{\mathbf{x}} \in S \text{ , whenever } \mathbf{x} \in S, \tilde{\mathbf{x}} \in S, \text{ and } \lambda \in [0,1]$$

  convex combination

- The interaction of convex sets is convex.
- Let f be a function of many variables defined on the convex set S. Then f is
  - concave on the set S if for all $\mathbf{x} \in S$, all $\tilde{\mathbf{x}} \in S$, and all $\lambda \in [(0,1]$ we have

$$f((1-\lambda)\mathbf{x}+\lambda\tilde{\mathbf{x}}) \geq (1-\lambda)f(\mathbf{x})+\lambda f(\tilde{\mathbf{x}})$$

  - convex on the set S if for all $\mathbf{x} \in S$, all $\tilde{\mathbf{x}} \in S$, and all $\lambda \in (0,1)$ we have

$$f((1-\lambda)\mathbf{x}+\lambda\tilde{\mathbf{x}}) \leq (1-\lambda)f(\mathbf{x})+\lambda f(\tilde{\mathbf{x}})$$

- f is concave $\Leftrightarrow$ $-f$ is convex

Calculus definition of concavity

- Linear restriction of function f: $g(t) = f(t\mathbf{x} + (1-t)\tilde{\mathbf{x}})$
- f is concave in $\mathbf{x} \Leftrightarrow g$ is concave in t

$$f(t\mathbf{x} + (1-t)\tilde{\mathbf{x}}) = g(t) = g(t \times 1 + (1-t) \times 0)$$

$$tf(\mathbf{x}) + (1-t)f(\tilde{\mathbf{x}}) = tg(1) + (1-t)g(0)$$

- g is concave in $t \Leftrightarrow g''(t) \leq 0$ or $\frac{g(t)-g(\tilde{t})}{(t-\tilde{t})} \leq g'(\tilde{t})$
- f is concave in $\mathbf{x} \Leftrightarrow f(\mathbf{x}) - f(\tilde{\mathbf{x}}) \leq \nabla f(\tilde{\mathbf{x}})'(\mathbf{x} - \tilde{\mathbf{x}})$

$$f(\mathbf{x}) - f(\tilde{\mathbf{x}}) = g(1) - g(0)$$

$$g'(0) = \nabla f(\tilde{\mathbf{x}})'(\mathbf{x} - \tilde{\mathbf{x}})$$

Conditions for concavity(convexity) and Hessian matrix

Let f be a function of many variables with continuous partial derivatives of first and second order on the convex open set S and denote the Hessian of f at the point $\mathbf{x}$ by $H_f(\mathbf{x})$. Then

- f is concave $\Leftrightarrow H_f(\mathbf{x})$ is negative semidefinite for all $\mathbf{x} \in S$
- if $H_f(\mathbf{x})$ is negative definite for all $x \in S \Rightarrow f$ is strictly concave
- f is convex $\Leftrightarrow H_f(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in S$
- if $H_f(\mathbf{x})$ is positive definite for all $\mathbf{x} \in S \Rightarrow f$ is strictly convex

Properties of concave function

- If the functions f and g are concave (convex) and $a \geq 0$ and $b \geq 0$ then the function $af + bg$ is concave (convex).
- If the function U is concave (convex) and the function g is nondecreasing and concave (convex) then the function f defined by $f(\mathbf{x}) = g(U(\mathbf{x}))$ is concave (convex).
  - Concavity is not necessarily preserved under monotone transformation $\Longrightarrow$ not an ordinal property.

$$\begin{aligned} f(t\mathbf{x} + (1-t)\tilde{\mathbf{x}}) &\geq tf(\mathbf{x}) + (1-t)f(\tilde{\mathbf{x}}) \\ g(f(t\mathbf{x} + (1-t)\tilde{\mathbf{x}})) &\geq g(tf(\mathbf{x}) + (1-t)f(\tilde{\mathbf{x}})) \end{aligned}$$

but no guarantee that

$$g(f(t\mathbf{x} + (1-t)\tilde{\mathbf{x}})) \geq tg(f(\mathbf{x})) + (1-t)g(f(\tilde{\mathbf{x}}))$$

More properties

Upper contour set of a concave function is convex everywhere

- Let f be a multivariate function defined on the convex set S. For any real a,
    - Upper level set of f for a: $P_a = \{\mathbf{x} \in S : f(\mathbf{x}) \geq a\}$
    - Lower level set of f for a: $P_a = \{\mathbf{x} \in S : f(\mathbf{x}) \leq a\}$

More properties

Upper contour set of a concave function is convex everywhere

- Let f be a multivariate function defined on the convex set S. For any real a,
  - Upper level set of f for a: $P_a = \{\mathbf{x} \in S : f(\mathbf{x}) \geq a\}$
  - Lower level set of f for a: $P_a = \{\mathbf{x} \in S : f(\mathbf{x}) \leq a\}$
- This feature can be preserved under monotone transformation $\Longrightarrow$ "*having convex upper contour sets*" is an ordinal property

$$
\begin{aligned}
& h(\mathbf{x}) = g(f(\mathbf{x})) \geq a, \; h(\tilde{\mathbf{x}}) = g(f(\tilde{\mathbf{x}})) \geq a, \; g \uparrow \\
\Rightarrow \quad & f(\mathbf{x}) \geq g^{-1}(a) \text{ and } f(\tilde{\mathbf{x}}) \geq g^{-1}(a) \\
\Rightarrow \quad & f((1-\lambda)\mathbf{x} + \lambda\tilde{\mathbf{x}}) \geq g^{-1}(a) \\
\Rightarrow \quad & h((1-\lambda)\mathbf{x} + \lambda\tilde{\mathbf{x}}) = g(f((1-\lambda)\mathbf{x} + \lambda\tilde{\mathbf{x}})) \geq g(g^{-1}(a)) = a
\end{aligned}
$$

Concavity and quasi-concavity

- **Concavity** is a **cardinal** property; **quasi-concavity** is an **ordinal** *"generalization"* of concavity.
- The multivariate function f defined on a convex set S is quasi-concave (quasi-convex) if every upper (lower) level set of f is convex.
- f is quasi-convex if and only if $-f$ is quasi-concave.

Alternative definition of quasi-concavity

- A function f is **quasi-concave** if and only if for all $\mathbf{x} \in S$, all $\tilde{\mathbf{x}} \in S$, and all $\lambda \in [0, 1]$ we have

$$\text{if } f(\mathbf{x}) \geq f(\tilde{\mathbf{x}}) \text{ then } f((1-\lambda)\mathbf{x} + \lambda\tilde{\mathbf{x}}) \geq f(\tilde{\mathbf{x}})$$

  a weaker condition than
  $f((1-\lambda)\boldsymbol{x} + \lambda\tilde{\boldsymbol{x}}) \geq (1-\lambda)f(\boldsymbol{x}) + \lambda f(\tilde{\boldsymbol{x}}))$
  as $(1-\lambda)f(\boldsymbol{x}) + \lambda f(\tilde{\boldsymbol{x}})) \geq f(\tilde{\boldsymbol{x}})$

- The multivariate function f defined on a convex set S is **strictly quasi-concave** if for all $\mathbf{x} \in S$, all $\tilde{\mathbf{x}} \in S$ with $\tilde{\mathbf{x}} \neq \mathbf{x}$, and all $\lambda \in (0, 1)$ we have

$$\text{if } f(\mathbf{x}) \geq f(\tilde{\mathbf{x}}) \text{ then } f((1-\lambda)\mathbf{x} + \lambda\tilde{\mathbf{x}}) > f(\tilde{\mathbf{x}})$$

Conditions for quasi-concavity

- $\nabla f(\tilde{\mathbf{x}})'(\mathbf{x} - \tilde{\mathbf{x}}) \geq 0$ if $f(\mathbf{x}) \geq f(\tilde{\mathbf{x}})$
  again, this is weaker than
  $\nabla f(\tilde{\boldsymbol{x}})'(\boldsymbol{x} - \tilde{\boldsymbol{x}}) \geq f(\boldsymbol{x}) - f(\tilde{\boldsymbol{x}})$ *if* $f(\boldsymbol{x}) \geq f(\tilde{\boldsymbol{x}})$

Conditions for quasi-concavity

- $\nabla f(\tilde{\mathbf{x}})'(\mathbf{x} - \tilde{\mathbf{x}}) \geq 0$ if $f(\mathbf{x}) \geq f(\tilde{\mathbf{x}})$
  again, this is weaker than
  $\nabla f(\tilde{\boldsymbol{x}})'(\boldsymbol{x} - \tilde{\boldsymbol{x}}) \geq f(\boldsymbol{x}) - f(\tilde{\boldsymbol{x}})$ *if* $f(\boldsymbol{x}) \geq f(\tilde{\boldsymbol{x}})$
- Bordered Hessian of $f(\mathbf{x})$, $H_f^B(\mathbf{x})$:

$$\begin{bmatrix} 0 & \nabla f(\mathbf{x})' \\ \nabla f(\mathbf{x}) & H_f(\mathbf{x}) \end{bmatrix}$$

Conditions for quasi-concavity,continue

- Let f be a function of n variables with continuous partial derivatives of first and second order in an open convex set S. Use D_k for the kth leading principal minor of $H_f^B(\mathbf{x})$.
  - If f is quasi-concave then the D_ks alternate in sign, starting with $D_3(\mathbf{x}) \geq 0$, $D_4(\mathbf{x}) \leq 0$,...,$D_n(\mathbf{x}) \geq 0$, and so on.
  - If D_ks alternate in sign, starting with $D_3(\mathbf{x}) > 0$, $D_4(\mathbf{x}) < 0$ and so on, then f is quasi-concave

Conditions for quasi-concavity,continue

- Let f be a function of n variables with continuous partial derivatives of first and second order in an open convex set S. Use D_k for the kth leading principal minor of $H_f^B(\mathbf{x})$.
  - If f is quasi-concave then the D_ks alternate in sign, starting with $D_3(\mathbf{x}) \geq 0$, $D_4(\mathbf{x}) \leq 0$,...,$D_n(\mathbf{x}) \geq 0$, and so on.
  - If D_ks alternate in sign, starting with $D_3(\mathbf{x}) > 0$, $D_4(\mathbf{x}) < 0$ and so on, then f is quasi-concave
  - If f is quasi-convex then $D_k(\mathbf{x}) \leq 0$ for all k, for all $\mathbf{x}$ in S
  - If $D_k(\mathbf{x}) < 0$ for all $k \geq 2$, for all $\mathbf{x}$ in S then f is quasi-convex

Optimization: set the stage

- Optimization Problem:

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

subject to $\mathbf{x} \in S$
where f is the **objective function**, $\mathbf{x}$ is the **choice variable**, S is the **constraint set**.

Optimization: set the stage

- Optimization Problem:

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

  subject to $\mathbf{x} \in S$
  where f is the **objective function**, $\mathbf{x}$ is the **choice variable**, S is the **constraint set**.
- The solution to the optimization problem $\mathbf{x}^*$ is called **maximizer**; $f(\mathbf{x}^*)$ is called the **maximum (value)** of the function f subject to the constraint $\mathbf{x} \in S$.

Optimization: set the stage

- Optimization Problem:

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

  subject to $\mathbf{x} \in S$
  where f is the **objective function**, $\mathbf{x}$ is the **choice variable**, S is the **constraint set**.
- The solution to the optimization problem $\mathbf{x}^*$ is called **maximizer**; $f(\mathbf{x}^*)$ is called the **maximum (value)** of the function f subject to the constraint $\mathbf{x} \in S$.
- Local maximizer: the variable $\mathbf{x}^*$ is a local maximizer of the function f subject to the constraint $\mathbf{x} \in S$ if there is a number $\epsilon > 0$ such that $f(\mathbf{x}) \leq f(\mathbf{x}^*)$ for all $\mathbf{x} \in S$ for which the distance between $\mathbf{x}$ and $\mathbf{x}^*$ is at most ϵ.

Optimization: set the stage

- Optimization Problem:

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

  subject to $\mathbf{x} \in S$
  where f is the **objective function**, $\mathbf{x}$ is the **choice variable**, S is the **constraint set**.
- The solution to the optimization problem $\mathbf{x}^*$ is called **maximizer**; $f(\mathbf{x}^*)$ is called the **maximum (value)** of the function f subject to the constraint $\mathbf{x} \in S$.
- Local maximizer: the variable $\mathbf{x}^*$ is a local maximizer of the function f subject to the constraint $\mathbf{x} \in S$ if there is a number $\epsilon > 0$ such that $f(\mathbf{x}) \leq f(\mathbf{x}^*)$ for all $\mathbf{x} \in S$ for which the distance between $\mathbf{x}$ and $\mathbf{x}^*$ is at most ϵ.
- Global maximizer

Existence of solution

- Bounded set: the set S is bounded if there exists a number k such that the distance of every point in S from the origin is at most k
- Compact set: a closed and bounded set is a compact set.
- A continuous function on a compact set attains both a maximum and a minimum on the set.

Interior optima

- Let f be a differentiable function of n variables defined on the set S. If the point $\mathbf{x}$ in the interior of S is a local or global maximizer or minimizer of f then

$$f_i(x) = 0 \text{ for } i = 1, ..., n$$

Interior optima

- Let f be a differentiable function of n variables defined on the set S. If the point $\mathbf{x}$ in the interior of S is a local or global maximizer or minimizer of f then

$$f_i(x) = 0 \text{ for } i = 1, ..., n$$

- Let f be a function of n variables with continuous partial derivatives of first and second order, defined on the set S. Suppose that $\mathbf{x}^*$ is a stationary point of f in the interior of S (so that $f_i(\mathbf{x}) = 0$ for all i).
  - If $H(\mathbf{x}^*)$ is negative definite then $\mathbf{x}^*$ is a local maximizer
  - If $\mathbf{x}^*$ is a local maximizer then $H(\mathbf{x}^*)$ is negative semidefinite
  - If $H(\mathbf{x}^*)$ is positive definite then $\mathbf{x}^*$ is a local minimizer
  - If $\mathbf{x}^*$ is a local minimizer then $H(\mathbf{x}^*)$ is positive semidefinite

Interior optima

- Let f be a differentiable function of n variables defined on the set S. If the point $\mathbf{x}$ in the interior of S is a local or global maximizer or minimizer of f then

$$f_i(x) = 0 \text{ for } i = 1, ..., n$$

- Let f be a function of n variables with continuous partial derivatives of first and second order, defined on the set S. Suppose that $\mathbf{x}^*$ is a stationary point of f in the interior of S (so that $f_i(\mathbf{x}) = 0$ for all i).
  - If $H(\mathbf{x}^*)$ is negative definite then $\mathbf{x}^*$ is a local maximizer
  - If $\mathbf{x}^*$ is a local maximizer then $H(\mathbf{x}^*)$ is negative semidefinite
  - If $H(\mathbf{x}^*)$ is positive definite then $\mathbf{x}^*$ is a local minimizer
  - If $\mathbf{x}^*$ is a local minimizer then $H(\mathbf{x}^*)$ is positive semidefinite
- Saddle point: a stationary point that is neither a local maximizer nor a local minimizer is a saddle point.

Interior optima, continue

Suppose that the function f has continuous partial derivatives in a convex set S and let $\mathbf{x}$ be in the interior of S. Then

- if f is concave
  then $\mathbf{x}$ is a global maximizer of f in $S \Leftrightarrow \mathbf{x}$ is a stationary point of f
- if f is convex
  then $\mathbf{x}$ is global minimizer of f in $S \Leftrightarrow \mathbf{x}$ is a stationary point of f

Equality constraints: necessary conditions

Let f and $g^1, ..., g^m$ be continuously differentiable functions of n variables defined on the set S, let c_j for $j = 1, ..., m$ be numbers, and suppose that $\mathbf{x}^*$ is an interior point of S that solves the problem

$$max_{\mathbf{x}} f(\mathbf{x}) \text{ subject to } g^j(\mathbf{x}) = c_j \text{ for } j = 1, ..., m$$

or the problem

$$min_{\mathbf{x}} f(\mathbf{x}) \text{ subject to } g^j(\mathbf{x}) = c_j \text{ for } j = 1, ..., m$$

or is a local maximizer or minimizer of $f(\mathbf{x})$ subject to $g^j(\mathbf{x}) = c_j$ for $j = 1, ..., m$.

Necessary conditions, continue

Suppose also that the rank of the Jacobian matrix J_g in which the (i,j)th component is $\frac{\partial g^j}{\partial x_i}|_{\mathbf{x}^*}$ is m (constraint qualification).
Then there are unique numbers $\lambda_1, ..., \lambda_m$ such that x^* is a stationary point of the Lagrangian function L defined by

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

That is, $\mathbf{x}^*$ statisfies the first-order conditons

$$L_i(\mathbf{x}^*, \boldsymbol{\lambda}) = f_i(\mathbf{x}^*) - \sum_{j=1}^{m} \lambda_j \frac{\partial g^j}{\partial x_i}|_{\mathbf{x}^*} = 0 \text{ for } i = 1, ..., n$$

In addition, $g^j(\mathbf{x}^*) = c_j$ for $j = 1, ..., m$.
In other words, we can solve

$$\nabla L(\mathbf{x}^*, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}^*) - J_g(\mathbf{x}^*)\boldsymbol{\lambda}^* = 0$$

for the $\boldsymbol{\lambda}$.

Second order conditions and bordered Hessian

Hessian of the equivalent unconstrained optimization problem

$$max_{\mathbf{x},\boldsymbol{\lambda}}\ L(\mathbf{x},\boldsymbol{\lambda})$$

where

$$L(\mathbf{x},\boldsymbol{\lambda}) = f(x_1,...,x_n) + \sum_{j=1}^{m} \lambda_j(c_j - g^j(x_1,...,x_n))$$

is the Hessian of f with a border. The bordered Hessian H^B is

$$\begin{bmatrix} \nabla_{\boldsymbol{\lambda}\boldsymbol{\lambda}} L(\mathbf{x},\boldsymbol{\lambda}) & \nabla_{\boldsymbol{\lambda}\mathbf{x}} L(\mathbf{x},\boldsymbol{\lambda}) \\ \nabla_{\mathbf{x}\boldsymbol{\lambda}} L(\mathbf{x},\boldsymbol{\lambda}) & \nabla_{\mathbf{x}\mathbf{x}} L(\mathbf{x},\boldsymbol{\lambda}) \end{bmatrix}$$

notice that $[\nabla_{\boldsymbol{\lambda}\boldsymbol{\lambda}} f(\mathbf{x},\boldsymbol{\lambda})] = \mathbf{0}$. Use $J_g(\mathbf{x})$ for the Jacobian matrix of the first order partial derivatives of the constraint functions g^js with respect to $\mathbf{x}$, H^B is then (after multiplying the last n rows and columns with -1)

$$\begin{bmatrix} \mathbf{0} & J_g(\mathbf{x})' \\ J_g(\mathbf{x}) & H_L(\mathbf{x}) \end{bmatrix}$$

Second order conditions,continue

Since on the left top corner is a $\mathbf{0}$ matrix, only the largest $n-m$ leading principal minors are non-trivial.
Suppose $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ statisfies the first-order conditons

$$L_i(\mathbf{x}^*, \boldsymbol{\lambda}^*) = f_i(\mathbf{x}^*) - \sum_{j=1}^{m} \lambda_j^* \frac{\partial g^j}{\partial x_i}|_{\mathbf{x}^*} = 0 \text{ for } i = 1, ..., n$$

When the g functions are linear in $\mathbf{x}$

- the largest $n-m$ leading principal minors have alternate signs with the largest having $(-1)^n$ at $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$
  $\Rightarrow$ FOCs give a local maximum
- the largest $n-m$ leading principal minors have the sign as $(-1)^m$ at $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$
  $\Rightarrow$ FOCs give a local minimum

Sufficient conditions

Suppose that f and g^j for $j = 1, ..., m$ are continuously differentiable functions defined on an open convex subset S of n-dimensional space and let $x^* \in S$ be an interior stationary point of the Lagrangian

$$L(\mathbf{x}) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j (g^j(\mathbf{x}) - c_j)$$

Suppose further that $g^j(\mathbf{x}^*) = c_j$ for $j = 1, ..., m$. Then

- if L is concave - in particular if f is concave and $\lambda_j^* g^j$ is convex for $j = 1, ..., m$ - then $\mathbf{x}^*$ solves the constrained maximization problem
- if L is convex - in particular if f is convex and $\lambda_j^* g^j$ is concave for $j = 1, ..., m$ - then $\mathbf{x}^*$ solves the constrained minimization problem

Lagrange method: constrained $\Rightarrow$ unconstrained

Transform the constrained problem into an unconstrained problem with m more variables

$$max_{\mathbf{x}}\ f(\mathbf{x}) \text{ subject to } g^j(\mathbf{x}) = c_j \text{ for } j = 1, ..., m$$

versus

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

- First and second order conditions
- Special cases where f is concave (convex) and the g^js are convex(concave)

Why Lagrange method works?

First order conditions

$$\frac{\partial L}{\partial x_i}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = f_i(\mathbf{x}^*) - \sum_{j=1}^{m} \lambda_j^* \frac{\partial g^j}{\partial x_i}|_{\mathbf{x}^*} = 0 \text{ for } i = 1, ..., n$$

$$\frac{\partial L}{\partial \lambda_j i}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = g^j(\mathbf{x}^*) - c_j = 0 \text{ for } j = 1, ..., m$$

A special case with only one constraint

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

Why Lagrange method works, continue

- $\nabla f(\mathbf{x}^*) = \lambda^* \nabla g(\mathbf{x}^*)$
  $\Rightarrow$ Directional derivative of f at $\mathbf{x}^*$ is proportional to directional derivative of g at $\mathbf{x}^*$
  $\Rightarrow$ When directional derivative of g is 0, the directional derivative of f is 0
- $g(\mathbf{x}) = c$
  $\Rightarrow$ $\mathbf{x}$ can only change in directions that gives 0 directional derivative of g; otherwise the constraint would be violated
- For constrained optimization problem, instead of setting gradient of the objective function to 0 (as in the unconstrained problem), we set the directional derivative to 0 with the direction of deviation specified by the constraint equation, i.e, we only care about deviations from $\mathbf{x}^*$ that satisfies

$$g_1(\mathbf{x}^*)\Delta x_1 + g_2(\mathbf{x}^*)\Delta x_2 +, \ldots, + g_n(\mathbf{x}^*)\Delta x_n = 0$$

- Generalization to multiple constraints $\Rightarrow$ more restrictions on the directional deviations $\mathbf{x}$ may take

Lagrange multiplier

- $\lambda^* = \frac{\nabla f(\mathbf{x}^*)'\Delta\mathbf{x}}{\nabla g(\mathbf{x}^*)'\Delta\mathbf{x}}$
- The rate of change in the maximum over a change in the constraint
- The value of the Lagrange multiplier on the jth constraint at the solution of the problem is equal to the rate of change in the maximal value of the objective function as the jth constraint is relaxed.
  Shadow price of scarce resource (internal value/imputed value)

Envelope theorem, unconstrained optimization

Let f be a continuously differentiable function of $n + k$ variables. Define the function f^* of k variables by

$$f^*(\mathbf{r}) = max_{\mathbf{x}} f(\mathbf{x}, \mathbf{r}) = f(\mathbf{x}^*(\mathbf{r}), \mathbf{r})$$

where $\mathbf{x}$ is an n-vector and $\mathbf{r}$ is a k-vector. If the solution of the maximization problem is a contiuuously differentiable function of $\mathbf{r}$ then

$$f_h^*(\mathbf{r}) = f_{n+h}(\mathbf{x}^*(\mathbf{r}), \mathbf{r})$$

Instead of the total derivative, you only need to check the partial derivative w.r.t. the parameter of interest at the optimizing point.

Envelope theorem, constrained optimization

Let f and g^j for $j = 1, ..., m$ be continuously differentiable functions of $n + k$ variables. Define the function f^* of k variables by

$$L^*(\mathbf{r}) = max_{\mathbf{x}} f(\mathbf{x}, \mathbf{r}) \text{ subject to } g^j(\mathbf{x}, \mathbf{r}) = 0 \text{ for } j = 1, ..., m$$

where $\mathbf{x}$ is an n-vector and $\mathbf{r}$ is a k-vector. Suppose that the solution of the maximization problem and the associated Lagrange multipliers $\lambda_1, ..., \lambda_m$ are continuously differentiable functions of r and the rank of the ("Jacobian") matrix in which the (i, j)th component is $\frac{\partial g^j}{\partial x_i}|_{\mathbf{x}^*}$ is m. Then

$$f_h^*(\mathbf{r}) = L_{n+h}(\mathbf{x}^*(\mathbf{r}), \mathbf{r}) \text{ for } h = 1, ..., k$$

where the function L is defined by
$L(\mathbf{x}, \mathbf{r}) = f(\mathbf{x}, \mathbf{r}) - \sum_{j=1}^m \lambda_j g^j(\mathbf{x}, \mathbf{r})$ for every $(\mathbf{x}, \mathbf{r})$

Inequality constraints

Optimization with inequality constraints

$$max_{\mathbf{x}} f(\mathbf{x}) \text{ subject to } g^j(\mathbf{x}) \leq c_j \text{ for } j = 1, ..., m$$

Lagrange function

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

Kuhn-Tucker conditions

The Kuhn-Tucker conditions for the problem are

$$(1) \quad L_i(\mathbf{x}, \boldsymbol{\lambda}) = f_i(\mathbf{x}^*) - \sum_{j=1}^{m} \lambda_j^* \frac{\partial g^j}{\partial x_i}|_{\mathbf{x}^*} = 0 \text{ for } i = 1, ..., n$$

$$(2) \quad g^j(\mathbf{x}) - c_j = 0 \ \& \ \lambda_j \geq 0$$

$$\text{or} \quad g^j(\mathbf{x}) - c_j < 0 \ \& \ \lambda_j = 0 \text{ for } j = 1, ..., m$$

The second set of conditions are equivalent to

$$\lambda_j \geq 0, g^j(\mathbf{x}) \leq c_j \text{ and } \lambda_j(g^j(\mathbf{x}) - c_j) = 0 \text{ for } j = 1, ..., m$$

Kuhn-Tucker conditions

The Kuhn-Tucker conditions for the problem are

$$(1) \quad L_i(\mathbf{x}, \boldsymbol{\lambda}) = f_i(\mathbf{x}^*) - \sum_{j=1}^{m} \lambda_j^* \frac{\partial g^j}{\partial x_i}|_{\mathbf{x}^*} = 0 \text{ for } i = 1, ..., n$$

$$(2) \quad g^j(\mathbf{x}) - c_j = 0 \ \& \ \lambda_j \geq 0$$

$$\text{or} \quad g^j(\mathbf{x}) - c_j < 0 \ \& \ \lambda_j = 0 \text{ for } j = 1, ..., m$$

The second set of conditions are equivalent to

$$\lambda_j \geq 0, g^j(\mathbf{x}) \leq c_j \text{ and } \lambda_j(g^j(\mathbf{x}) - c_j) = 0 \text{ for } j = 1, ..., m$$

- *One of two inequalities must be binding - "complementary slackness condition"*
- *If a Lagrange multiplier is positive, its associated constraint must be binding; if a constraint is slack, its associated Lagrange multiplier must be zero.*

Necessity of the Kuhn-Tucker conditions

Let f and g^j for $j = 1, ..., m$ be continuously differentiable functions of many variables and let c_j for $j = 1, ..., m$ be constants. Suppose that $\mathbf{x}^*$ solves the problem

$$max_{\mathbf{x}}\, f(\mathbf{x}) \text{ subject to } g^j(\mathbf{x}) \leq c_j \text{ for } j = 1, ..., m$$

Suppose that (constraint qualification)

- either the gradient vectors $\nabla g^j(\mathbf{x}^*)$ associated with constraint j that binds at $\mathbf{x}^*$ are linearly independent
- or each g^j is concave
- or each g^j is convex and there is some $\mathbf{x}$ such that $g^j(\mathbf{x}) < c_j$ for $j = 1, ..., m$
- or each g^j is quasi-convex, $\nabla g^j(\mathbf{x}^*) \neq \mathbf{0}$ for all j, and there is some x such that $g^j(\mathbf{x}) < c_j$ for $j = 1, ..., m$

Then there exists a unique vector $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_m)'$ such that $(\mathbf{x}^*, \boldsymbol{\lambda})$ satisfies the Kuhn-Tucker conditions.

Sufficiency of the Kuhn-Tucker conditions

Let f and g^j for $j = 1, ..., m$ be continuously differentiable functions of many variables and let c_j for $j = 1, ..., m$ be constants. Consider the problem

$$max_{\mathbf{x}}\, f(\mathbf{x}) \text{ subject to } g^j(\mathbf{x}) \leq c_j \text{ for } j = 1, ..., m$$

Suppose that

- f is concave
- and g^j is quasi-convex for $j = 1, ..., m$

If there exists $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_m)'$ such that $(\mathbf{x}^*, \boldsymbol{\lambda})$ satisfies the Kuhn-Tucker conditions then $\mathbf{x}^*$ solves the problem

Necessary and sufficient

The Kuhn-Tucker conditions are both necessary and sufficient if

- the objective function is concave
- and
  - either each constraint is linear
  - or each constraint function is convex and some vector of the variables satisfies all constraints strictly

Weaker sufficient conditions

Let f and g^j for $j = 1, ..., m$ be continuously differentiable functions of many variables and let c_j for $j = 1, ..., m$ be constants. Consider the problem

$$max_{\mathbf{x}} f(\mathbf{x}) \text{ subject to } g^j(\mathbf{x}) \leq c_j \text{ for } j = 1, ..., m$$

Suppose that

- f is twice-differentiable and quasi-concave
- and g^j is quasi-convex for $j = 1, ..., m$

If there exists $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_m)'$ such that $(\mathbf{x}^*, \boldsymbol{\lambda})$ satisfies the Kuhn-Tucker conditions and $f_i(\mathbf{x}^*) \neq 0$ for some $i = 1, ..., n$ then $\mathbf{x}^*$ solves the problem

Necessary and sufficient

Suppose that

- the objective function is twice differentiable and quasi-concave
- every constraint is linear

Then

- if $\mathbf{x}^*$ solves the problem then there exists a unique vector $\boldsymbol{\lambda}$ such that $(\mathbf{x}^*, \boldsymbol{\lambda})$ satisfies the Kuhn-Tucker conditions
- if $(\mathbf{x}^*, \boldsymbol{\lambda})$ satisfies the Kuhn-Tucker conditions and $f_i(\mathbf{x}^*) \neq 0$ for some $i = 1, ..., n$ then $\mathbf{x}^*$ solves the problem.