# Calculus

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differentiable functions</td>
<td>2</td>
</tr>
<tr>
<td>Derivation of necessary conditions for a maximum</td>
<td>3</td>
</tr>
<tr>
<td>Sufficient conditions</td>
<td>6</td>
</tr>
<tr>
<td>Contour sets and tangent hyperplanes</td>
<td>10</td>
</tr>
<tr>
<td>Derivation of necessary conditions for constrained maximization problems</td>
<td>15</td>
</tr>
<tr>
<td>Integrals</td>
<td>25</td>
</tr>
<tr>
<td>Paradox</td>
<td>53</td>
</tr>
</tbody>
</table>
Differentiable functions

$X = [a, b]$

A function $f : X \to \mathbb{R}$ is differentiable at $x^0$ if

$$\lim_{x \to x^0} \frac{f(x) - f(x^0)}{x - x^0}$$

has a limit at $x^0$.

This limit is called the derivative of the function at $x^0$

$$\frac{df}{dx}(x^0) \equiv f'(x^0) \equiv Df(x^0) = \lim_{x \to x^0} \frac{f(x) - f(x^0)}{x - x^0}$$

A function is differentiable on $X$ if it is differentiable for all $x \in X$.

A function is continuously differentiable on $X$ if the mapping $\frac{df}{dx} : X \to \mathbb{R}$ is continuous.

($f \in C^1$ on $X$.)
Formal derivation of necessary conditions for a maximum

Lemma 1: If \( \frac{df}{dx}(x^0) > 0 \) then for some \( \delta \) – neighborhood, \( N(x^0, \delta) \),

\[ x > x^0 \Rightarrow f(x) > f(x^0) \quad \text{and} \quad x < x^0 \Rightarrow f(x) < f(x^0) \]

Proof: If \( \frac{df}{dx}(x^0) > 0 \) then for any \( \varepsilon > 0 \) there exists a deleted \( \delta \) – neighborhood, \( N^D(x^0, \delta) \) such that for all \( x \) in this deleted neighborhood

\[
\frac{df}{dx}(x^0) - \varepsilon < \frac{f(x) - f(x^0)}{x - x^0} < \frac{df}{dx}(x^0) + \varepsilon
\]

Choose \( \varepsilon = \frac{1}{2} \frac{df}{dx}(x^0) \). Then

\[
0 < \frac{1}{2} \frac{df}{dx}(x^0) < \frac{f(x) - f(x^0)}{x - x^0} \quad \text{for any} \quad x \quad \text{in} \quad N^D(x^0, \delta). \quad \text{QED}
\]

By an almost identical argument we also have

Lemma 2: If \( \frac{df}{dx}(x^0) < 0 \) then for some \( \delta \) – neighborhood, \( N(x^0, \delta) \),

\[ x > x^0 \Rightarrow f(x) < f(x^0) \quad \text{and} \quad x < x^0 \Rightarrow f(x) < f(x^0) \]
Necessary conditions for a maximum

\[ x \in \mathbb{R} \quad \text{Max} \ f(x) \text{ where } X = [\alpha, \beta] \]

For an interior max \( x^0 \in \text{int} \ X = (\alpha, \beta) \)

**First Order Conditions (FOC)**

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

**Proof by contradiction:**

Suppose \( \frac{df}{dx}(x^0) > 0 \). By Lemma 1 for some \( \delta \)–neighborhood, \( N(x^0, \delta) \),

if \( x > x^0 \) then \( f(x) > f(x^0) \)

But then \( f(x) \) cannot take on its maximum at \( x^0 \) after all.

By Lemma 2 if \( x < x^0 \) then \( f(x) > f(x^0) \)

QED
Second Order Conditions (SOC)

\[ \frac{d^2 f}{dx^2}(x^0) \leq 0 \]

Proof by contradiction:

Suppose that \( x^0 \in \text{int} \ X \)

Suppose \( \frac{d}{dx} \frac{df}{dx}(x^0) > 0 \). By Lemma 1, \( \frac{df}{dx}(x) - \frac{df}{dx}(x^0) > 0 \), in some deleted \( \delta \) – neighborhood of \( x^0 \).

For any \( \hat{x} \in (x^0, x^0 + \delta) \) it follows that \( \frac{df}{dx}(\hat{x}) - \frac{df}{dx}(x^0) > 0 \). Appealing to the FOC, \( \frac{df}{dx}(x^0) = 0 \) it follows that \( \frac{df}{dx}(\hat{x}) > 0 \). Then, by Lemma 1 \( f(x) > f(x^0) \) for all \( x > x^0 \) in \( N^D(x^0, \delta) \).

Class exercise:

What are the necessary condition for a maximum on the boundary of \( X \)?
Sufficient conditions for a local maximum

\[
\frac{df}{dx}(x^0) = 0, \quad \frac{d}{dx} \left( \frac{df}{dx}(x^0) \right) < 0
\]

By Lemma 2, in some deleted delta neighborhood of \( x^0 \)

\[
\frac{df}{dx}(x) - \frac{df}{dx}(x^0) = \frac{df}{dx}(x) < 0.
\]

Therefore for some \( \delta > 0 \) \( \frac{df}{dx}(x) < 0 \) on \( (x^0 - \delta, x^0 + \delta) \). Therefore \( \frac{df}{dx}(x) > 0 \) on \( (x^0 - \delta, x^0) \) and on \( (x^0, x^0 + \delta) \) and \( \frac{df}{dx}(x) < 0 \) on \( (x^0 - \delta, x^0) \).

By Lemma 1 \( f(x) \) is strictly increasing on \( (x^0 - \delta, x^0) \).

By Lemma 2 \( f(x) \) is strictly decreasing on \( (x^0, x^0 + \delta) \).

QED
A set $X$ is convex if for any pair of points in the set the line segment connecting these points also lies in the set.

$$f: X \rightarrow \mathbb{R} \text{ where } X \text{ is a convex subset of } \mathbb{R}^n$$

**Maximization**

$$\text{Max } f(x)_{x \in X}$$

Necessary conditions for $f$ to take on its maximum at $x^0$

Define

$$g(\lambda) = f(x(\lambda)) \text{ where } x(\lambda) = x^0 + \lambda(x^1 - x^0)$$

$g(\cdot)$ is a mapping from an interval $[a,b]$ into $\mathbb{R}$

Thus we can appeal to our previous results.

$$\frac{dg}{d\lambda}(\lambda) = \frac{\partial f}{\partial x}(x(\lambda)) \cdot (x^1 - x^0)$$

**FOC (interior max)**

$$\frac{\partial f}{\partial x}(x^0) = 0$$
SOC

\[
\frac{dg}{d\lambda} (\lambda) = \frac{\partial f}{\partial x}(x(\lambda)) \cdot (x^1 - x^0) = \frac{\partial f}{\partial x}(x(\lambda))'(x^1 - x^0) = (x^1 - x^0)' \frac{\partial f}{\partial x}(x(\lambda))
\]

\[
\frac{d}{d\lambda} \frac{dg}{d\lambda} (\lambda) =
\begin{bmatrix}
\frac{d}{d\lambda} \frac{\partial f}{\partial x_i}(x(\lambda)) \\
\vdots \\
\frac{d}{d\lambda} \frac{\partial f}{\partial x_n}(x(\lambda))
\end{bmatrix} \cdot
\begin{bmatrix}
\frac{\partial}{\partial x} \frac{\partial f}{\partial x_i}(x(\lambda)) \\
\vdots \\
\frac{\partial}{\partial x} \frac{\partial f}{\partial x_n}(x(\lambda))
\end{bmatrix}
= (x^1 - x^0)' \cdot (x^1 - x^0) = (x^1 - x^0)' \left[ \frac{\partial^2 f}{\partial x_i \partial x_j} \right] (x^1 - x^0)
\]

SOC  The “Hessian Matrix” of second partial derivatives must be negative semi definite.
Sufficient Conditions

If the Hessian matrix is negative definite then \( f \) has a local maximum at \( x^0 \)

Exercise:

If the FOC and SOC both hold at \( x^0 \) does this imply that the function takes on its maximum at \( x^0 \)?
Contour sets

\[ f : S \subset \mathbb{R}^n \to \mathbb{R} \]

Assume that \( f \) is continuously differentiable (\( f \in C^1 \)) that is, the partial derivatives are all continuous

**Contour set for** \( x^0 \in S \)

\[ C(x^0) = \{ x \in S \mid f(x) = f(x^0) \} \]

Example: indifference curve
**Contour sets**

\[ f : S \subset \mathbb{R}^n \rightarrow \mathbb{R} \]

Assume that \( f \) is continuously differentiable (\( f \in C^1 \)) that is, the partial derivatives are all continuous

**Contour set for** \( x^0 \in S \)

\[ C(x^0) = \{ x \in S \mid f(x) = f(x^0) \} \]

Example: indifference curve

**Upper contour set for** \( x^0 \in S \)

\[ C^U(x^0) = \{ x \in S \mid f(x) \geq f(x^0) \} \]

Example: consumption vectors preferred to \( x^0 \)
Contour sets

\[ f : S \subset \mathbb{R}^n \rightarrow \mathbb{R} \]

Assume that \( f \) is continuously differentiable (\( f \in \mathcal{C}^1 \)) that is, the partial derivatives are all continuous

**Contour set for** \( x^0 \in S \)

\[ C(x^0) = \{ x \in S \mid f(x) = f(x^0) \} \]

Example: indifference curve

**Upper contour set for** \( x^0 \in S \)

\[ C^U(x^0) = \{ x \in S \mid f(x) \geq f(x^0) \} \]

Example: consumption vectors preferred to \( x^0 \)

**Lower contour set for** \( x^0 \in S \)

\[ C^L(x^0) = \{ x \in S \mid f(x) \leq f(x^0) \} \]

Example:

Budget set \( f(x) = p \cdot x \),

\[ C^L(x^0) = \{ x \in \mathbb{R}^n_+ \mid p \cdot x \leq p \cdot x^0 \} \]
**Tangent line**

Consider the line

\[ T(x^0) = \{ x | \frac{\partial f}{\partial x_1}(x^0)(x_1 - x_1^0) + \frac{\partial f}{\partial x_2}(x^0)(x_2 - x_2^0) \} \]

The slope of this line is

\[ -\frac{\frac{\partial f}{\partial x_1}(x^0)}{\frac{\partial f}{\partial x_2}(x^0)} \]
**Tangent plane**

Consider the line

\[ T(x^0) = \{ x | \frac{\partial f}{\partial x_1}(x^0)(x_1 - x_1^0) + \frac{\partial f}{\partial x_2}(x^0)(x_2 - x_2^0) \} \]

The slope of this line is

\[ -\frac{\frac{\partial f}{\partial x_1}(x^0)}{\frac{\partial f}{\partial x_2}(x^0)} \]

Now consider the contour set \( C(x^0) = \{ x | f(x) = f(x^0) \} \) at a point where \( \frac{\partial f}{\partial x_2}(x^0) \neq 0 \).
Tangent plane

Consider the line

\[ T(x^0) = \{ x \mid \frac{\partial f}{\partial x_1}(x^0)(x_1 - x_1^0) + \frac{\partial f}{\partial x_2}(x^0)(x_2 - x_2^0) \} \]

The slope of this line is

\[ \frac{\partial f}{\partial x_1}(x^0) - \frac{\partial f}{\partial x_2}(x^0) \]

Now consider the contour set \( C(x^0) = \{ x \mid f(x) = f(x^0) \} \)

at a point where \( \frac{\partial f}{\partial x_2}(x^0) \neq 0 \).

The contour set (at least locally) implicitly defines the function \( x_2 = h(x_1) \).

Along the contour set \( \frac{d}{dx_1} f(x_1, h(x_1)) = 0 \)

Applying the Chain Rule for differentiating a function of a function

\[ \frac{d}{dx_1} f(x_1, h(x_1)) = \frac{\partial f}{\partial x_1}(x) + \frac{\partial f}{\partial x_2}(x) \frac{dh}{dx_1} = 0 \]
It follows that the slope of the contour set at \( x^0 \) is \( \frac{dh}{dx_1}(x^0) = -\frac{\partial f}{\partial x_1}(x^0) \).

Because the line and curve have the same slope at \( x^0 \) the line is called the tangent line.

**Exercise:** \( f(x) = x_1^2 + x_2^2 \). What is the tangent line (i) if \( x^0 = (3,4) \) (ii) \( x^0 = (0,5) \) (iii) \( x^0 = (5,0) \)
Generalization to $\mathbb{R}^n$.

**Definition: Tangent hyperplane at $x^0$**

$$T(x^0) = \{x \mid \frac{\partial f}{\partial x}(x^0) \cdot (x - x^0) = 0\}$$
Generalization to \( \mathbb{R}^n \).

**Definition: Tangent hyperplane at** \( x^0 \)

\[
T(x^0) = \{ x \mid \frac{\partial f}{\partial x}(x^0) \cdot (x - x^0) = 0 \}
\]

Choose some point \( x^1 \) in the set

\[
\{ x \mid \frac{\partial f}{\partial x}(x^0) \cdot (x - x^0) > 0 \}
\]

In the figure \( x^1 \) lies to the right of the tangent line.
Generalization to \( \mathbb{R}^n \).

**Definition: Tangent hyperplane at** \( x^0 \)

\[
T(x^0) = \{ x \mid \frac{\partial f}{\partial x}(x^0) \cdot (x - x^0) = 0 \}
\]

Choose some point \( x^1 \) in the set

\[
\{ x \mid \frac{\partial f}{\partial x}(x^0) \cdot (x - x^0) > 0 \}
\]

In the figure \( x^1 \) lies to the right of the tangent line.

Consider the points on the line segment connecting \( x^0 \) and \( x^1 \).

These are the convex combinations \( x(\lambda) = x^0 + \lambda(x^1 - x^0) \)

We will argue that for sufficiently small \( \lambda \), \( f(x(\lambda)) > f(x^0) \).
Tangent Hyperplane Lemma: Suppose that $f(x) \in \mathbb{C}^1$ and that $x^1$ lies in the interior of the set $X = \{x \mid \frac{\partial f}{\partial x}(x^0) \cdot (x - x^0) \geq 0\}$. Then for $\lambda \in (0,1)$ and sufficiently small, $f(x(\lambda)) > f(x^0)$, where $x(\lambda) = (1 - \lambda)x^0 + \lambda x^1$.

Proof:

Define $g(\lambda) = f(x(\lambda))$

where $x(\lambda) = (1 - \lambda)x^0 + \lambda x^1 = x^0 + \lambda(x^1 - x^0)$.

Note that $\frac{dx_j}{d\lambda} = x_j^1 - x_j^0$ for $i = 1, \ldots, n$. 
Proof:

Define \( g(\lambda) \equiv f(x(\lambda)) \)

where \( x(\lambda) = (1 - \lambda)x^0 + \lambda x^1 = x^0 + \lambda(x^1 - x^0) \).

Note that \( \frac{dx_j}{d\lambda} = x_j^1 - x_j^0 \) for \( i = 1, ..., n \) .

Apply the Chain Rule

\[
\frac{dg}{d\lambda} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x(\lambda)) \frac{dx_j}{d\lambda}
\]
Proof:

Define $g(\lambda) \equiv f(x(\lambda))$

where $x(\lambda) = (1 - \lambda)x^0 + \lambda x^1 = x^0 + \lambda(x^1 - x^0)$.

Note that $\frac{dx_j}{d\lambda} = x^1_j - x^0_j \quad i = 1, \ldots, n$.

Apply the Chain Rule

$$\frac{dg}{d\lambda} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x(\lambda)) \frac{dx_j}{d\lambda}$$

Therefore

$$\frac{dg}{d\lambda}(\lambda) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x(\lambda)) (x_j^1 - x_j^0) = \frac{\partial f}{\partial x}(x(\lambda)) \cdot (x^1 - x^0).$$
Proof:

Define \( g(\lambda) \equiv f(x(\lambda)) \)

where \( x(\lambda) = (1 - \lambda)x^0 + \lambda x^1 = x^0 + \lambda(x^1 - x^0) \).

Note that \( \frac{dx_j}{d\lambda} = x^1_j - x^0_j \quad i = 1, \ldots, n \).

Apply the Chain Rule

\[
\frac{dg}{d\lambda} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x(\lambda)) \frac{dx_j}{d\lambda}
\]

Therefore

\[
\frac{dg}{d\lambda}(\lambda) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x(\lambda))(x^1_j - x^0_j) = \frac{\partial f}{\partial x}(x(\lambda)) \cdot (x^1 - x^0).
\]

Hence \( \frac{dg}{d\lambda}(0) = \frac{\partial f}{\partial x}(x^0) \cdot (x^1 - x^0) > 0 \).
Proof:

Define $g(\lambda) \equiv f(x(\lambda))$

where $x(\lambda) = (1 - \lambda)x^0 + \lambda x^1 = x^0 + \lambda(x^1 - x^0)$.

Note that $\frac{dx_j}{d\lambda} = x^1_j - x^0_j$ for $i = 1, \ldots, n$.

Apply the Chain Rule

$$\frac{dg}{d\lambda} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x(\lambda)) \frac{dx_j}{d\lambda}$$

Therefore

$$\frac{dg}{d\lambda}(\lambda) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x(\lambda))(x^1_j - x^0_j) = \frac{\partial f}{\partial x}(x(\lambda)) \cdot (x^1 - x^0).$$

Hence $\frac{dg}{d\lambda}(0) = \frac{\partial f}{\partial x}(x^0) \cdot (x^1 - x^0) > 0$.

Then, since the partial derivatives are continuous, over some interval $[0, \hat{\lambda}]$, $\frac{dg}{d\lambda}(\lambda) > 0$

Therefore over this interval $g(\lambda) = f(x(\lambda))$ is strictly increasing.

Q.E.D
**Constrained Optimization with linear constraints**

As a stepping stone to a formal derivation of necessary conditions for the general non-linear problem, consider the following problem with linear constraints

\[
P: \max_{x} \{ f(x) \mid x \geq 0, \ a_i \cdot x \geq 0, \ i = 1, \ldots, m \}. \ f \in \mathbb{C}^l
\]

Suppose the maximum is at \( \bar{x} \) and \( \frac{\partial f}{\partial x}(\bar{x}) \neq 0 \). Rather than solve \( P \) consider the linearized approximation

\[
P': \max_{x} \left\{ \frac{\partial f}{\partial x}(\bar{x}) \cdot x \mid x \geq 0, \ a_i \cdot x \geq 0, \ i = 1, \ldots, m \right\}. \ f \in \mathbb{C}^l,
\]

We now argue that if \( \bar{x} \) solves \( P \), then it must also solve \( P' \).
**Proposition:** If $\bar{x}$ solves $P$: $\max_x \{ f(x) \mid x \geq 0, \ a_i \cdot x \geq 0, \ i = 1,\ldots,m \}$. $f \in \mathbb{C}^1$ and $\frac{\partial f}{\partial x}(\bar{x}) \neq 0$.

Then $\bar{x}$ must also solve $P'$: $\max_x \{ \frac{\partial f}{\partial x}(\bar{x}) \cdot x \mid x \geq 0, \ a_i \cdot x \geq 0, \ i = 1,\ldots,m \}$. $f \in \mathbb{C}^1$.

**Proof:** (by contradiction)

Suppose that $\bar{x}$ that solves $P$ but not $P'$. Then there is some $\hat{x}$ satisfying

$$a_i \cdot \hat{x} \geq 0, \ i = 1,\ldots,m$$

(*)

such that

$$\frac{\partial f}{\partial x}(\bar{x}) \cdot (\hat{x} - \bar{x}) = \frac{\partial f}{\partial x}(\bar{x}) \cdot \hat{x} - \frac{\partial f}{\partial x}(\bar{x}) \cdot \bar{x} > 0$$

(**)
**Proposition:** If $\bar{x}$ solves $P: \max \{ f(x) \mid x \geq 0, \ a_i \cdot x \geq 0, \ i = 1, \ldots, m \}$. $f \in \mathbb{C}^1$ and $\frac{\partial f}{\partial x}(\bar{x}) \neq 0$

Then $\bar{x}$ must also solve $P': \max \{ \frac{\partial f}{\partial x}(\bar{x}) \cdot x \mid x \geq 0, \ a_i \cdot x \geq 0, \ i = 1, \ldots, m \}$. $f \in \mathbb{C}^1$.

**Proof:** (by contradiction)

Suppose that $\bar{x}$ that solves $P$ but not $P'$. Then there is some $\hat{x}$ satisfying

$$a_i \cdot \hat{x} \geq 0, \ i = 1, \ldots, m$$

such that

$$\frac{\partial f}{\partial x}(\bar{x}) \cdot (\hat{x} - \bar{x}) = \frac{\partial f}{\partial x}(\bar{x}) \cdot \hat{x} - \frac{\partial f}{\partial x}(\bar{x}) \cdot \bar{x} > 0$$

Consider the convex combination $x^\lambda$ of $\bar{x}$ and $\bar{x}$. Since

$$a_i \cdot \hat{x} \geq 0, \ i = 1, \ldots, m \text{ and } a_i \cdot \bar{x} \geq 0, \ i = 1, \ldots, m$$

it follows that $x^\lambda = (1 - \lambda)\bar{x} + \lambda \hat{x}$ is feasible for all $\lambda \in (0,1)$.

Appealing to the Tangent Hyperplane Lemma, for sufficiently small $\lambda$, $f(x^\lambda) > f(\bar{x})$.

But then $\bar{x}$ is not a maximizer for $P$.

QED
Necessary conditions

It follows that any necessary conditions for $\bar{x}$ to solve the linear problem $P'$ are also necessary conditions for $\bar{x}$ to solve the problem $P$.

While the proof is a little more complicated, the next proposition establishes the conditions under which this approach works for the general non-linear problem.

These conditions are very mild restrictions on the constraints (and are known as the constraint qualifications).

**Exercise: Necessary conditions when the feasible set is convex.**

Show that if $\bar{x}$ solves $P$: $\max_x f(x) \mid x \in X$ where $f \in \mathbb{C}^1$ and $X$ is a convex set, then $\bar{x}$ must also solve $P'$: $\max_x \frac{\partial f(\bar{x})}{\partial x} \cdot x \mid x \geq 0, \ x \in X$. $f \in \mathbb{C}^1$. 

**Constrained** maximization for the general non-linear problem

Problem: \( \max_x \{ f(x) \mid x \geq 0, h_i(x) \geq 0, \ i = 1,\ldots,m \} \). \( f \in \mathbb{R}^l, \ h_i \in \mathbb{R}^l, \ i = 1,\ldots,m \)

Suppose the maximum is at \( \bar{x} \) and that \( \frac{\partial f}{\partial x}(\bar{x}) \neq 0 \).

Suppose \( k \) constraints are binding at \( \bar{x} \). Re-label so that they are the first \( k \) constraints. Then

\( \bar{x} \in \arg \max_x \{ f(x) \mid x \geq 0, h_i(x) \geq 0,\ldots,h_k(x) \geq 0 \} \).
Constrained maximization for the general non-linear problem

Problem: \( \max_x \{ f(x) \mid x \geq 0, h_i(x) \geq 0, \ i = 1,\ldots,m \} \). \( f \in \mathbb{C}^1, \ h_i \in \mathbb{C}^1, \ i = 1,\ldots,m \)

Suppose the maximum is at \( \bar{x} \) and that \( \frac{\partial f}{\partial x}(\bar{x}) \neq 0 \).

Suppose \( k \) constraints are binding at \( \bar{x} \). Re-label so that they are the first \( k \) constraints. Then
\[
\bar{x} \in \text{arg} \max_x \{ f(x) \mid x \geq 0, h_1(x) \geq 0,\ldots,h_k(x) \geq 0 \}.
\]

Consider the linearized maximization problem
\[
\max_x \{ \frac{\partial f}{\partial x}(\bar{x}) \cdot (x - \bar{x}) \mid x \geq 0, \frac{\partial h_i}{\partial x}(\bar{x}) \cdot (x - \bar{x}) \geq 0, \ i = 1,\ldots,k \}.
\]

For there to be \( k \) constraints we require that \( \frac{\partial h_i}{\partial x}(\bar{x}) \neq 0, \ i = 1,\ldots,k \).
Constrained maximization for the general non-linear problem

Problem: \( \max_x \{ f(x) \mid x \geq 0, h_i(x) \geq 0, \ i = 1, \ldots, m \} \). \( f \in \mathbb{C}^l, \ h_i \in \mathbb{C}^l, \ i = 1, \ldots, m \)

Suppose the maximum is at \( \bar{x} \) and that \( \frac{\partial f}{\partial x}(\bar{x}) \neq 0 \).

Suppose \( k \) constraints are binding at \( \bar{x} \). Re-label so that they are the first \( k \) constraints. Then
\[
\bar{x} \in \arg \max_{x} \{ f(x) \mid x \geq 0, h_i(x) \geq 0, i = 1, \ldots, k \}.
\]

Consider the linearized maximization problem
\[
\max_x \{ \frac{\partial f}{\partial x}(\bar{x}) \cdot (x - \bar{x}) \mid x \geq 0, \frac{\partial h_i}{\partial x}(\bar{x}) \cdot (x - \bar{x}) \geq 0, \ i = 1, \ldots, k \}.
\]

For there to be \( k \) constraints we require that \( \frac{\partial h_i}{\partial x}(\bar{x}) \neq 0, \ i = 1, \ldots, k \).

Suppose finally that the linearized feasible set has a non-empty interior.

Then
\[
\bar{x} \in \arg \max_x \{ \frac{\partial f}{\partial x}(\bar{x}) \cdot (x - \bar{x}) \mid x \geq 0, \frac{\partial h_i}{\partial x}(\bar{x}) \cdot (x - \bar{x}) \geq 0, i = 1, \ldots, k \}
\]

Therefore to obtain necessary conditions \( \bar{x} \) to be the solution of the nonlinear problem we look for necessary conditions for \( \bar{x} \) to solve this linear maximization problem.
\[
\hat{x} = (1 - \mu)\hat{x} + \mu x^0
\]

\[
x^\lambda = (1 - \lambda)\bar{x} + \lambda\hat{x}
\]

\[
\frac{\partial h}{\partial x}(\bar{x}) \cdot (x - \bar{x}) = 0
\]

\[
\frac{\partial h}{\partial x}(\bar{x}) \cdot (x - \bar{x}) \geq 0
\]
\[ f(x) = f(\bar{x}) \]

\[ \frac{\partial f}{\partial x}(\bar{x}) \cdot (x - \bar{x}) \geq 0 \]

\[ x^0 = (1 - \lambda)\bar{x} + \lambda\hat{x} \]

\[ \hat{x} = (1 - \mu)\hat{x} + \mu x^0 \]
As we have already seen, necessary conditions for $\bar{x}$ to solve the linear maximization problem are obtained by appealing to the supporting hyperplane theorem. These are the Kuhn-Tucker conditions.
Integrals:

**Indefinite integral**

\[ F(x) \text{ is the indefinite integral of } f(x) \text{ if } f(x) \text{ is the derivative of } F(x) \]
Integrals:

Indefinite integral

\[ F(x) \] is the indefinite integral of \( f(x) \) if \( f(x) \) is the derivative of \( F(x) \).

Application: Social Benefit

Typically measured as area under the demand price function \( p(q) \).

Note that \( p(\hat{q})\Delta q \geq \Delta B \geq p(\hat{q} + \Delta q)\Delta q \)

Therefore \( p(\hat{q}) \geq \frac{\Delta B}{\Delta q} \geq p(\hat{q} + \Delta q) \).

Taking the limit, \( \frac{dB}{dq} = p(q) \).
Integrals:

Indefinite integral

\( F(x) \) is the indefinite integral of \( f(x) \) if \( f(x) \) is the derivative of \( F(x) \)

Application: Social Benefit

Typically measured as area under the demand price function \( p(q) \).

Note that \( p(\hat{q}) \Delta q \geq \Delta B \geq p(\hat{q} + \Delta q) \Delta q \)

Therefore \( p(\hat{q}) \geq \frac{\Delta B}{\Delta q} \geq p(\hat{q} + \Delta q) \).

Taking the limit, \( \frac{dB}{dq} = p(q) \).

Integrals:

Indefinite Integral

If \( p(q) = a - bq \) then the indefinite integral is

\[
\int p(q)\,dq = aq - \frac{1}{2}bq^2 + K
\]

for any constant \( K \)

The indefinite integral is usually written without the constant term.
**Definite Integral:**

Value of increasing output from $q_1$ to $q_2$

$$B(q_2) = \int_{q_1}^{q_2} p(q) dq$$

If $p(q) = a - bq$ the indefinite integral is

$$aq - \frac{1}{2} bq^2$$
**Definite Integral:**

Value of increasing output from \( q_1 \) to \( q_2 \)

\[
B(q_2) = \int_{q_1}^{q_2} p(q) \, dq
\]

If \( p(q) = a - bq \) the indefinite integral is

\[
aq - \frac{1}{2}bq^2
\]

For some \( K \),

\[
B(q_2) = aq - bq^2 + K
\]

Since \( B(q_1) = 0 \) it follows that

\[
K = -(aq_1 - \frac{1}{2}bq_1^2)
\]

Therefore

\[
B(q_2) = \int_{q_1}^{q_2} p(q) \, dq = (aq_2 - \frac{1}{2}bq_2^2) - (aq_1 - \frac{1}{2}bq_1^2)
\]

Note that this is the indefinite integral evaluated at \( q_2 \) less the indefinite integral evaluated at \( q_1 \).
Integration by parts

Define $f''(x) = \frac{df}{dx}$ and $G(x) = \int g(z)dz$

Appealing to the Chain Rule

$$\frac{d}{dx} fG = f'(x)G(x) + f(x)g(x)$$
Integration by parts

Define \( f''(x) = \frac{df}{dx} \) and \( G(x) = \int g(z)dz \)

Appealing to the Chain Rule

\[
\frac{d}{dx} fG = f'(x)G(x) + f(x)g(x)
\]

Reintegrating,

\[
fG = \int f'(x)G(x)dx + \int f(x)g(x)dx
\]
Integration by parts

Define \( f'(x) = \frac{df}{dx} \) and \( G(x) = \int g(z)dz \)

Appealing to the Chain Rule

\[
\frac{d}{dx} fG = f'(x)G(x) + f(x)g(x)
\]

Reintegrating,

\[
fG = \int f'(x)G(x)dx + \int f(x)g(x)dx
\]

Rearranging

\[
\int f(x)g(x)dx = fG - \int f'(x)G(x)dx
\]

Integral of \( f \times g = f \times \text{integral of } g - \text{integral of } [\text{derivative of } f \times \text{integral of } g] \)
Example

Integrate ∫ ln x dx.

Use a little cunning.

∫ ln x dx = ∫ (ln x)(1) dx

Define f(x) = ln x, g(x) = 1. Then f′(x) = \frac{1}{x} and G(x) = x

∫ f(x)g(x) = fG − ∫ f′(x)G(x) = (ln x)x − ∫ \frac{1}{x} x dx = x ln x - x

Check by differentiating.
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma) \, dx \]

At what rate does \( G \) vary with \( \alpha \)?
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma) dx \]

At what rate does \( G \) vary with \( \alpha \)?

\[ G(\alpha + h, \beta, \gamma) = \int_{\alpha+h}^{\beta} f(x, \gamma) dx \]

\[ G(\alpha + h, \beta, \gamma) - G(\alpha, \beta, \gamma) \approx -f(x + h, \gamma)h \]
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma) \, dx \]

At what rate does \( G \) vary with \( \alpha \)?

\[ G(\alpha + h, \beta, \gamma) = \int_{\alpha + h}^{\beta} f(x, \gamma) \, dx \]

\[ G(\alpha + h, \beta, \gamma) - G(\alpha, \beta, \gamma) \approx -f(x + h, \gamma)h \]

\[ \frac{G(\alpha + h, \beta, \gamma) - G(\alpha, \beta, \gamma)}{h} \approx -f(x + h, \gamma) \]
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma) \, dx \]

At what rate does \( G \) vary with \( \alpha \)?

\[ G(\alpha + h, \beta, \gamma) = \int_{\alpha+h}^{\beta} f(x, \gamma) \, dx \]

\[ G(\alpha + h, \beta, \gamma) - G(\alpha, \beta, \gamma) \approx -f(x + h, \gamma)h \]

\[ \frac{G(\alpha + h, \beta, \gamma) - G(\alpha, \beta, \gamma)}{h} \approx -f(x + h, \gamma) \]

Take the limit as \( h \to 0 \) and so \( \delta x \to 0 \).

\[ \frac{\partial G}{\partial \alpha} = -f(\alpha, \gamma) \]
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma) \, dx \]

At what rate does \( G \) vary with \( \alpha \)?

\[ G(\alpha + h, \beta, \gamma) = \int_{\alpha + h}^{\beta} f(x, \gamma) \, dx \]

\[ G(\alpha + h, \beta, \gamma) - G(\alpha, \beta, \gamma) \approx -f(x + h, \gamma)h \]

\[ \frac{G(\alpha + h, \beta, \gamma) - G(\alpha, \beta, \gamma)}{h} \approx -f(x + h, \gamma) \]

Take the limit as \( h \to 0 \) and so \( \delta x \to 0 \).

\[ \frac{\partial G}{\partial \alpha} = -f(\alpha, \gamma) \]

By an almost identical argument

\[ \frac{\partial G}{\partial \beta} = f(\beta, \gamma) \]
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma) \, dx \]

This is the limit of a sum of rectangles of width \( \delta x \)

\[ S_n(\alpha, \beta, \gamma) = \sum_{i=1}^{n} f(x_i, \gamma) \delta x \]
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma)\,dx \]

This is the limit of a sum of rectangles of width $\delta x$

\[ S_n(\alpha, \beta, \gamma) = \sum_{t=1}^{n} f(x_t, \gamma)\delta x \]

\[ S_n(\alpha, \beta, \gamma + h) - S_n(\alpha, \beta, \gamma) = \sum_{t=1}^{n} f(x_t, \gamma + h)\delta x - \sum_{t=1}^{n} f(x_t, \gamma)\delta x \]

\[ = \sum_{t=1}^{n} [f(x_t, \gamma + h) - f(x_t, \gamma)]\delta x \]
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma) \, dx \]

This is the limit of a sum of rectangles of width \( \delta x \)

\[ S_n(\alpha, \beta, \gamma) = \sum_{i=1}^{n} f(x_i, \gamma) \delta x \]

\[ S_n(\alpha, \beta, \gamma + h) - S_n(\alpha, \beta, \gamma) = \sum_{i=1}^{n} f(x_i, \gamma + h) \delta x - \sum_{i=1}^{n} f(x_i, \gamma) \delta x \]

\[ = \sum_{i=1}^{n} [f(x_i, \gamma + h) - f(x_i, \gamma)] \delta x \]

\[ \frac{S_n(\alpha, \beta, \gamma + h) - S_n(\alpha, \beta, \gamma)}{h} = \sum_{i=1}^{n} \left[ \frac{f(x_i, \gamma + h) - f(x_i, \gamma)}{h} \right] \delta x \]

Take the limit with respect to \( h \)

\[ \frac{\partial S_n}{\partial \gamma} = \sum_{i=1}^{n} \frac{\partial f}{\partial \gamma}(x_i, \gamma) \delta x \]
Consider the definite integral

\[ G(\alpha, \beta, \gamma) = \int_{\alpha}^{\beta} f(x, \gamma) \, dx \]

This is the limit of a sum of rectangles of width \( \delta x \)

\[ S_n(\alpha, \beta, \gamma) = \sum_{t=1}^{n} f(x_t, \gamma) \delta x \]

\[ S_n(\alpha, \beta, \gamma + h) - S_n(\alpha, \beta, \gamma) = \sum_{t=1}^{n} f(x_t, \gamma + h) \delta x - \sum_{t=1}^{n} f(x_t, \gamma) \delta x \]

\[ = \sum_{t=1}^{n} [f(x_t, \gamma + h) - f(x_t, \gamma)] \delta x \]

\[ \frac{S_n(\alpha, \beta, \gamma + h) - S(\alpha, \beta, \gamma)}{h} = \sum_{t=1}^{n} \left[ \frac{f(x_t, \gamma + h) - f(x_t, \gamma)}{h} \right] \delta x \]

Take the limit with respect to \( h \)

\[ \frac{\partial S_n}{\partial \gamma} = \sum_{t=1}^{n} \frac{\partial f}{\partial \gamma}(x_t, \gamma) \delta x \]

Take the limit with respect to \( n \)

\[ \frac{\partial G}{\partial \gamma} = \int_{\alpha}^{\beta} \frac{\partial f}{\partial \gamma}(x, \gamma) \, dx \]
**Paradox**

A function is everywhere differentiable on \( \mathbb{R} \) and takes on its maximum at \( x^0 \). However there may be no neighborhood \( N(x^0, \delta) \) over which the derivative \( \frac{df}{dx}(x) \) is decreasing.

If this seems paradoxical to you it is probably because it is tempting to think that a differentiable function must have a slope that varies continuously. However, as the following example illustrates, this is not necessarily true.

\[
f(x) = \begin{cases} 
0, & x = 0 \\
-x^2(2 + \sin(1/x)), & x \neq 0.
\end{cases}
\]

Differentiating,

\[
\frac{df}{dx}(x) = -2x(1 + \sin(1/x)) + \cos(1/x), \ x \neq 0
\]

The cosine varies from -1 to +1 with every \( 2\pi \) radians thus as \( x \) approaches zero and so \( 1/x \) increases ever more rapidly, the second term fluctuates from -1 to +1 ever more rapidly while the first term approaches zero. Thus the derivative does not converge.
\[ f(x) = \begin{cases} 0, & x = 0 \\ -x^2(2 + \sin(1/x)), & x \neq 0 \end{cases} \]

Note next that \(-3x^2 \leq f(x) \leq -x^2\). Thus \(f(x)\) has its maximum at \(x = 0\).

Moreover

\[-3x^2 \leq f(x) - f(0) \leq -x^2.\]

Therefore

\[-\frac{3x^2}{x} \leq \frac{f(x) - f(0)}{x - 0} \leq -\frac{x^2}{x}.\]

Since the upper and lower bounds approach zero it follows that the ratio approaches zero in the limit.

Therefore \(\frac{df}{dx}(0) = 0\).
The graph of the function is as depicted.

Also shown are the upper and lower bounds.

While the slope of the function oscillates ever more rapidly between -1 and 1 as $x \to 0$,

the derivative is well defined at zero.
Exercise (for those who like mathematical technicalities.)

Consider the function $f(x) = x(2 + \sin(1/x)), \ x \neq 0$ and $f(0) = 0.$

(a) Show that the function is continuous.
(b) Show that the function is strictly increasing at $x = 0.$
(c) Show that the slope is unbounded from below and above in any delta-neighborhood of $x = 0.$