Optimization Problem

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $D \subset \mathbb{R}^n$. A constrained optimization problem is

$$\max f(x) \quad \text{subject to } x \in D$$

- $f$ is the objective function
- $D$ is the constraint set
- A solution to this problem is $x \in D$ such that
  $$f(x) \geq f(y) \quad \text{for all } y \in D$$

Such an $x$ is called a maximizer
- The set of maximizers is denoted
  $$\text{argmax}\{f(x) | x \in D\}$$

Similarly for minimization problems

Example

$$\max x^3 - 3x^2 + 2x + 1 \quad \text{subject to } 0.1 \leq x \leq 2.5$$
Existence of Solutions

**Theorem (Weierstrass Theorem)**

Let $D \subset \mathbb{R}^n$ be compact and $f : D \rightarrow \mathbb{R}$ be a continuous function on $D$. Then $f$ attains a maximum and a minimum on $D$.

**Theorem**

A set $D \subset \mathbb{R}^n$ is **compact** if and only if it is closed and bounded.

**Definition**

A set $D \subset \mathbb{R}^n$ is **bounded** if there exists $k > 0$ such that $\|x\| < k$ for each $x \in D$.

Here $\|x\| = (\sum_{i=1}^{n} x_i^2)^{1/2}$ is the Euclidean norm. If $x$ is a real number, then $\|x\|$ is simply its absolute value $|x|$.

**Definition**

A set $D \subset \mathbb{R}^n$ is **closed** if and only if for all sequences $(x_k)$ such that $x_k \in D$ for each $k$ and $x_k \rightarrow x$, we have $x \in D$.

The **Euclidean distance** between $x, y \in \mathbb{R}^n$ is given by

$$d(x, y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}$$

**Definition**

A sequence $(x_k)$ in $\mathbb{R}^n$ is said to converge to a limit $x$ (written $x_k \rightarrow x$) if for all $\epsilon > 0$, there exists an integer $K(\epsilon)$ such that for all $k \geq K(\epsilon)$, we have $d(x_k, x) < \epsilon$. 

---

**Example**

Max $-(x - 1)^2 + 2$ s.t. $x \in [0, 2]$.

---

**Example**

$D = [0, 2]$. For any $x \in [0, 2]$, $\|x\| = (x^2)^{1/2} = x$. Therefore, we can take $k = 3$, in which case $\|x\| < k$, and hence this set is bounded.

**Example**

$D = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq 1\}$. Take any $x, y \in D$ and note that $y^2 \leq 1 - x^2$ implies that $y^2 \leq 1$ and hence $|y| \leq 1$. Similarly, $|x| \leq 1$. Therefore, $x^2 + y^2 = (|x|)^2 + (|y|)^2 \leq 2$. Therefore, $(x^2 + y^2)^{1/2} \leq \sqrt{2}$. Therefore, this set is bounded.
Example

Consider the sequence \((x_k)\), where \(x_k = \frac{1}{k}, k = 1, 2, \ldots\). This sequence converges to 0. Take any \(\varepsilon > 0\) and let \(K(\varepsilon)\) be the smallest integer larger than \(1/\varepsilon\). Then, \(k \geq K(\varepsilon)\) implies that \(d(x_k, 0) = 1/k \leq 1/K(\varepsilon) < \varepsilon\).

Example

\(D = [0, 1]\) is closed. Indeed, take any sequence \((x_k) \to x\), such that \(x_k \in [0, 1]\) for all \(k\). This implies that for any \(\varepsilon > 0\), there exists an integer \(K(\varepsilon)\) such that for all \(k \geq K(\varepsilon)\), we have \(d(x_k, x) = |x - x_k| < \varepsilon\), or \(x_k - \varepsilon < x < x_k + \varepsilon\). Since \(x_k \geq 0\) and \(x_k \leq 1\), this implies that for any \(\varepsilon > 0\), \(-\varepsilon < x < 1 + \varepsilon\). Since, \(\varepsilon\) is arbitrary this implies that \(0 \leq x \leq 1\).

Example

\(D = [0, 1]\) is not closed. Take the sequence \((x_k)\), where \(x_k = 1 - 1/k, k = 1, 2, \ldots\), and note that \(x_k \in [0, 1]\) for all \(k\). Also, \((x_k) \to 1\), but \(1 \notin [0, 1]\).

Definition

A function \(f : S \to T\) where \(S \subset \mathbb{R}^n\) and \(T \subset \mathbb{R}^l\) is continuous at \(x \in S\) if for all \(\varepsilon > 0\), there is a \(\delta > 0\) such that \(y \in S\) and \(d(x, y) < \delta\) implies \(d(f(x), f(y)) < \varepsilon\). Equivalently, \(f\) is continuous at \(x \in S\) if for all \((x_k)\) such that \(x_k \in S\) for each \(k\) and \(x_k \to x\), we have \(f(x_k) \to f(x)\). \(f\) is continuous on \(S\) if it is continuous at each \(x \in S\).

Definition

\(f : S \to \mathbb{R}\), where \(S\) is an open subset of \(\mathbb{R}\), is differentiable at \(x \in S\) if

\[
\lim_{h \to 0} \frac{f(x + h) - f(x)}{h}
\]

exists, in which case this limit is the derivative of \(f\) at \(x\), denoted \(f'(x)\) or \(Df(x)\).
If $Df$ is differentiable, i.e., each $f_i: S \rightarrow \mathbb{R}$ is differentiable, at $x$ we denote partial derivatives of $f_i$ with respect to $i$ and $j \neq i$ as

$$\frac{\partial^2 f(x)}{\partial x_i^2} \quad \text{and} \quad \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

We say that $f$ is twice differentiable at $x$, with second derivative $D^2 f(x)$ given by

$$D^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

If a function is (twice) continuously differentiable we say that it is $C^1$ ($C^2$). For a $C^2$ function

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$$

Example

Let $f(x) = -(x-1)^2 + 2$. Then, $Df(x) = f'(x) = -2(x-1)$ and $D^2 f(x) = f''(x) = -2$.

Example

Let $f(x) = x^2 - y + xy$. Then, $Df(x) = [2x+y, -1+x]$ and

$$D^2 f(x) = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$$

A Simple Case

Let $f: \mathbb{R} \rightarrow \mathbb{R}$ and consider the problem $\max_{x \in [a,b]} f(x)$.

We call a point $x^*$ such that $f'(x^*) = 0$ a critical point.

Interior Optima

Theorem

Let $f: \mathbb{R} \rightarrow \mathbb{R}$ and suppose $a < x^* < b$ is a local maximum (minimum) of $f$ on $[a,b]$. Then, $f'(x^*) = 0$.

- Known as first order conditions
- Only necessary for interior local optima
  - Not necessary for global optima
  - Not sufficient for local optima.
- To distinguish between interior local maximum and minimum you can use second order conditions

Theorem

Let $f: \mathbb{R} \rightarrow \mathbb{R}$ and suppose $a < x^* < b$ is a local maximum (minimum) of $f$ on $[a,b]$. Then, $f''(x^*) \leq 0$ ($f''(x^*) \geq 0$).
Recipe for solving the simple case

Let \( f : \mathbb{R} \rightarrow \mathbb{R} \) be a differentiable function and consider the problem \( \max_{x \in [a,b]} f(x) \). If the problem has a solution, then it can be found by the following method:

1. Find all critical points: i.e., \( x^* \in [a,b] \) s.t. \( f'(x^*) = 0 \)
2. Evaluate \( f \) at all critical points and at boundaries \( a \) and \( b \)
3. The one that gives the highest \( f \) is the solution

We can use Weierstrass theorem to determine if there is a solution

Note that if \( f'(a) > 0 \) (or \( f'(b) < 0 \)), then the solution cannot be at \( a \) (or \( b \))

Example

\[
\max_{x \in [-1, 2]} - (x - 1)^2 + 2 \text{ s.t. } x \in [0, 2].
\]

Solution

\( f \) is continuous and \([-1, 2]\) is closed and bounded, and hence compact. Therefore, by Weierstrass theorem the problem has a solution. \( f'(x) = -2(x - 1) = 0 \) is solved at \( x = 1 \), which is the only critical point. We have \( f(0) = 2, f(-1) = 1, f(2) = 1 \). Therefore, \( 1 \) is the global maximum.

Recipe for general problems

Generalizes to \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) and the problem is

\[
\max f(x) \quad \text{subject to } x \in D
\]

- Find critical points \( x^* \in D \) such that \( Df(x^*) = 0 \)
- Evaluate \( f \) at the critical points and the boundaries of \( D \)
- Choose the one that give the highest \( f \)

Important to remember that solution must exist for this method to work

In more complicated problems evaluating \( f \) at the boundaries could be difficult

For such cases we have the method of the Lagrangian (for equality constraints) and Kuhn-Tucker conditions (for inequality constraints)
Equality Constraints: Lagrangean Method

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ for each $i = 1, \ldots, k$ be $C^1$ and consider the problem

$$\max f(x) \text{ s.t. } g_i(x) = 0, \quad i = 1, \ldots, k$$

1. Form the Lagrangean (the $\lambda_i$’s are called Lagrange multipliers)

$$L(x, \lambda) = f(x) + \sum_{i=1}^{k} \lambda_i g_i(x)$$

2. Find critical points of $L(x, \lambda)$

$$\frac{\partial L}{\partial x_j}(x, \lambda) = 0, \quad j = 1, \ldots, n$$

$$\frac{\partial L}{\partial \lambda_i}(x, \lambda) = 0, \quad i = 1, \ldots, k$$

3. Evaluate $f$ at each critical point $x$, choose the one that gives the highest value.

---

**Proposition**

Suppose that

1. A global optimum $x^*$ exists
2. Constraint qualification is met at $x^*$

Then, there exists $\lambda^*$ such that $(x^*, \lambda^*)$ is a critical point of $L$.

- This result implies that if these conditions hold, then the Lagrangean method will identify the optimum.
- Constraint qualification is the condition that rank of the matrix $Dg(x^*)$ (the $ij$th element is $\partial g_i(x^*)/\partial x_j$) is equal to $k$. In a problem with two variables $x_1, x_2$, this is equivalent to $\partial g_i(x^*)/\partial x_1 \neq 0$ or $\partial g_i(x^*)/\partial x_2 \neq 0$.
- In many cases both conditions can be verified to hold before hand and the method can be safely applied.

---

**Example**

Maximize $f(x, y) = x^2 - y^2$ subject to $g(x, y) = 1 - x^2 - y^2 = 0$. 

**Solution**

1. Since $f$ is continuous and the constraint set is compact there is a global maximum.
2. $Dg(x, y) = (-2x, -2y)$ and $0^2 + 0^2 = 0 \neq 1$ implies that $Dg(x^*, y^*) \neq (0, 0)$ at any critical point $x^*$. Therefore, the constraint qualification is satisfied.

Now we can apply the method of Lagrangean:

$$L(x, y, \lambda) = x^2 - y^2 + \lambda(1 - x^2 - y^2)$$

2. Find the critical points:

$$\frac{\partial L(x)}{\partial x} = 2x - 2\lambda x = 0$$

$$\frac{\partial L(x)}{\partial y} = -2y - 2\lambda y = 0$$

$$\frac{\partial L(x)}{\partial \lambda} = 1 - x^2 - y^2 = 0$$

$\lambda = 1$ or $\lambda = -1$, otherwise first two equations imply $x = y = 0$, which contradicts the third equation. Using this fact it is easy to show that there are four possible solutions:

$$(x, y, \lambda) \in \{(1, 0, 1), (-1, 0, 1), (0, 1, -1), (0, -1, -1)\}$$
Solution (cont’d)

3. Evaluate \( f \) at the critical points:

\[
f(1, 0) = f(-1, 0) = 1 \quad f(0, 1) = f(0, -1) = -1
\]

We conclude that the first two points are global maximizers.

Inequality Constraints: Kuhn-Tucker Method

Let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) and \( h_i : \mathbb{R}^n \rightarrow \mathbb{R} \) for each \( i = 1, \ldots, l \) be \( C^1 \) and consider the problem

\[
\max f(x) \text{ s.t. } h_i(x) \geq 0, \quad i = 1, \ldots, l
\]

1. Form the Lagrangean (\( \lambda \)'s are called Lagrange multipliers)

\[
L(x, \lambda) = f(x) + \sum_{i=1}^{k} \lambda_i h_i(x)
\]

2. Find critical points of \( L(x, \lambda) \)

\[
\frac{\partial L}{\partial x_j}(x, \lambda) = 0, \quad j = 1, \ldots, n
\]

\[
\frac{\partial L}{\partial \lambda_i}(x, \lambda) \geq 0, \quad \lambda_i \geq 0, \quad \lambda_i \frac{\partial L}{\partial \lambda_i}(x, \lambda) = 0, \quad i = 1, \ldots, k
\]

3. Evaluate \( f \) at each critical point \( x \), choose the one that gives the highest value

Example

\[
\max f(x, y) = x^2 - y \quad \text{subject to } g(x, y) = 1 - x^2 - y^2 \geq 0.
\]

Solution

1. Since \( f \) is continuous and the constraint set is compact there is a global maximum.

2. When the constraint binds \( x^2 + y^2 = 1 \), and hence \( x \neq 0 \) or \( y \neq 0 \). This implies that \( Dg(x^*, y^*) \neq (0, 0) \) at any critical point \( x^* \) for which the constraint binds. Therefore, the constraint qualification is satisfied.

Now we can apply the method of Lagrangean.
Solution (cont’d)

1. Set up the Lagrangean:

\[ L(x, y, \lambda) = x^2 - y + \lambda(1 - x^2 - y^2) \]

2. Find the critical points:

\[
\begin{align*}
2x - 2\lambda x &= 0 \\
-1 - 2\lambda y &= 0 \\
1 - x^2 - y^2 &\geq 0, \quad \lambda \geq 0, \quad \lambda(1 - x^2 - y^2) = 0
\end{align*}
\]

From the first equation \( \lambda = 1 \) or \( x = 0 \). If \( \lambda = 1 \), then from the second equation \( y = -1/2 \) and from the third condition \( x^2 + y^2 = 1 \). This gives two critical points \((\pm\sqrt{3}/2, -1/2, 1)\). If \( x = 0 \), then from the second equation \( \lambda > 0 \). Then, we must have \( x^2 + y^2 = 1 \), and hence \( y = \pm 1 \). But \( y = 1 \) contradicts second equation and \( \lambda \geq 0 \), and hence the only possible critical value is \((0, -1, 1/2)\).

Solution (cont’d)

3. Evaluate \( f \) at the critical points:

\[ f(\sqrt{3}/2, -1/2) = 5/3 \quad f(0, -1) = 1 \]

We conclude that the first two points are global maximizers.

---

Convexity and Optimization

Under suitable convexity assumptions necessary conditions for local optima become sufficient for global optima:

- For maximization: concave objective function and convex constraint set
- For minimization: convex objective function and convex constraint set

Definition

A set \( D \subset \mathbb{R}^n \) is **convex** if for all \( x, y \in D \) and \( \lambda \in (0, 1) \), it is the case that \( \lambda x + (1 - \lambda)y \in D \).

---

Concave and Convex Functions

Let \( D \subset \mathbb{R}^n \) and \( f : D \to \mathbb{R} \). From now on we assume \( D \) is convex.

Definition

A function \( f : D \to \mathbb{R} \) is **concave** if for any \( x, y \in D \) and \( \lambda \in (0, 1) \), it is the case that

\[ f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y) \]

and **convex** if

\[ f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \]
Concave and Convex Functions

Definition
A function \( f : \mathcal{D} \to \mathbb{R} \) is **strictly concave** if for any \( x \neq y \in \mathcal{D} \) and \( \lambda \in (0, 1) \), it is the case that
\[
f(\lambda x + (1 - \lambda)y) > \lambda f(x) + (1 - \lambda)f(y)
\]
and **strictly convex** if
\[
f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)
\]

Theorem
\( f : \mathcal{D} \to \mathbb{R} \) is (strictly) concave if and only if \( -f \) is (strictly) convex.

Concave and Convex Functions

Theorem
Let \( f : \mathcal{D} \to \mathbb{R} \) be a \( C^2 \) function, where \( \mathcal{D} \subset \mathbb{R}^n \) is open and convex. Then,
1. \( f \) is concave if and only if \( D^2 f(x) \) is negative semidefinite for all \( x \in \mathcal{D} \).
2. \( f \) is convex if and only if \( D^2 f(x) \) is positive semidefinite for all \( x \in \mathcal{D} \).
3. If \( D^2 f(x) \) is negative definite for all \( x \in \mathcal{D} \), then \( f \) is strictly concave.
4. If \( D^2 f(x) \) is positive definite for all \( x \in \mathcal{D} \), then \( f \) is strictly convex.

Concave and Convex Functions

Example
Let \( f(x) = -(x-1)^2 + 2 \). Then, \( f''(x) = -2 < 0 \) and hence this function is strictly concave.

Concave and Convex Functions

Example
Let \( f(x) = 1 - x^2 - y^2 \). Then, \( f_{11} = -2 < 0 \) and \( f_{11}f_{22} - f_{12}^2 = (-2)(-2) - 0 > 0 \), and hence this function is strictly concave.
Convexity and Optimization

**Theorem**

Let \( f : D \to \mathbb{R} \) be concave and \( D \) convex. Then,
1. Any local maximum of \( f \) is a global maximum.
2. The set of maximizers is either empty or convex. If \( f \) is strictly concave, then the set of maximizers is either empty or contains a single point.

**Theorem**

Let \( f : D \to \mathbb{R} \) be differentiable and concave and \( D \) convex. Then, \( x \) is an interior maximum of \( f \) on \( D \) if and only if \( D f(x) = 0 \).

Convexity and Kuhn-Tucker

- \( U \subset \mathbb{R}^n \) open and convex
- \( f : U \to \mathbb{R} \) and \( h_i : U \to \mathbb{R} \) for each \( i = 1, \ldots, l \) concave and \( C^1 \)
- Consider the problem
  \[
  \max f(x) \text{ s.t. } h_i(x) \geq 0, \quad i = 1, \ldots, l
  \]
  - If there is some \( \bar{x} \in U \) such that \( h_i(\bar{x}) > 0 \) for all \( i = 1, \ldots, l \)
    - known as Slater’s condition
  - Kuhn-Tucker method will always identify global optima
  - No need to verify
    - existence of a solution
    - constraint qualification

Example

Let \( f : \mathbb{R} \to \mathbb{R} \) and consider the problem:

\[
\max f(x) = -(x - 2)^2 \text{ subject to } h(x) = x - 1 \geq 0
\]

Solution

1. \( f'(x) = -2(x - 2) \) and \( f''(x) = -2 < 0 \), and hence \( f \) is (strictly) concave; \( h''(x) = 0 \), and hence \( h \) is concave.
2. There is an \( x \in \mathbb{R} \) (e.g., \( x = 2 \)) such that \( h(x) > 0 \)

This implies

1. If the Lagrangean has no critical point, then there is no solution to the problem.
2. If \( (x^*, \lambda^*) \) is a critical point, then it is a solution. In fact, since \( f \) strictly concave, it is the unique solution.
Solution (cont’d)

1. Set up the Lagrangean:

\[ L(x, \lambda) = -(x - 2)^2 + \lambda(x - 1) \]

2. Find the critical points:

\[-2(x - 2) + \lambda = 0\]

\[ x - 1 \geq 0, \quad \lambda \geq 0, \quad \lambda(x - 1) = 0 \]

- First equation and \( \lambda \geq 0 \Rightarrow x \geq 2 \)
- \( \lambda(x - 1) = 0 \Rightarrow \lambda = 0 \)
- First equation \( \Rightarrow x = 2 \)

Therefore, the only critical point is \((2, 0)\), and \( x = 2 \) is the unique global maximizer.