1. Convex Closure of a Set
2. Epigraphs and Hypographs of Convex Functions
3. Separation of Convex Sets
4. Gâteaux and Directional Derivatives
5. Convex Functions
6. Differentiability and Convexity
If $S \subseteq \mathbb{R}^n$ and $\mathcal{C}_S$ is the collection of convex sets that contain $S$, then

1. $\mathcal{C}_S \neq \emptyset$ because $\mathbb{R}^n$ is a convex set that contains $S$.
2. Any intersection of subsets of $\mathcal{C}_S$ is a convex set that contains $S$.

Thus $\bigcap \mathcal{C}_S$ is the least convex set that contains $S$. 

Definition

The convex closure of the subset $S$ of $\mathbb{R}^n$ is the set $K_{\text{conv}}(S) = \bigcap C_S$.

The convex closure of $S$ is denoted by $K_{\text{conv}}(S)$. 

Note that

- $S \subseteq K_{\text{conv}}(S)$;
- $S_1 \subseteq S_2$ implies $K_{\text{conv}}(S_1) \subseteq K_{\text{conv}}(S_2)$;
- $K_{\text{conv}}(K_{\text{conv}}(S)) = K_{\text{conv}}(S)$. 

Definition

Let \( f : \mathbb{R}^n \rightarrow \hat{\mathbb{R}} \) be a function. Its \textit{epigraph} is the set

\[
\text{epi}(f) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq y\}.
\]

The \textit{hypograph} of \( f \) is the set

\[
\text{hyp}(f) = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid y \leq f(x)\}.
\]
The epigraph of a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is the dotted area in $\mathbb{R}^2$ located above the graph of the function $f$ and it is shown in Figure 1(a); the hypograph of $f$ is the dotted area below the graph shown in Figure 1(b).

**Figure**: Epigraph (a) and hypograph (b) of a function $f : \mathbb{R} \longrightarrow \mathbb{R}$
Note that the intersection

\[ \text{epi}(f) \cap \text{hyp}(f) = \{(x, y) \in S \times \mathbb{R} \mid y = f(x)\} \]

is the graph of the function \(f\).

If \(f(x) = \infty\), then \((x, \infty) \not\in \text{epi}(f)\). Thus, for the function \(f_{\infty}\) defined by \(f_{\infty}(x) = \infty\) for \(x \in S\) we have \(\text{epi}(f_{\infty}) = \emptyset\).
Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and let $a \in \mathbb{R}$. The \textit{level set for f at a} is the set

$$L_{f,a} = \{x \in S \mid f(x) \leq a\}.$$
Definition

Let $C, D$ be two subsets of $\mathbb{R}^n$. A hyperplane $w'x - a = 0$ separates $C, D$ if $w'x \leq a$ for every $x \in C$ and $w'x \geq a$ for every $x \in D$. If a separating hyperplane $H$ exists for two subsets $C, D$ of $L$ we say that $C, D$ are separable.

$C$ and $D$ are separated by a hyperplane $H$ if $C$ and $D$ are located in distinct closed half-spaces associated to $H$. The sets $C$ and $D$ are linearly separable if there exists a hyperplane that separates them.
Definition

The subsets $C$ and $D$ of $\mathbb{R}^n$ are strictly separated by a hyperplane $w'x = a$ if we have either $w'x > a > w'y$ for $x \in C$ and $y \in D$, or $w'y > a > w'x$ for $x \in C$ and $y \in D$.

The sets $C$ and $D$ are strictly linearly separable if there exists a hyperplane that strictly separates them.
**Theorem**

Let $C$ be a convex subset in $\mathbb{R}^n$ such that $\mathbf{l}(C) \neq \emptyset$ and let $V$ be an affine subspace such that $V \cap \mathbf{l}(C) = \emptyset$.

There exists a closed hyperplane $H$, $\mathbf{w}'x = a$, in $\mathbb{R}^n$ such that

1. $V \subseteq H$ and $H \cap \mathbf{l}(C) = \emptyset$, and
2. there exists $c \in \mathbb{R}$ such that $\mathbf{w}'x = c$ for all $x \in V$ and $\mathbf{w}'x < c$ for all $x \in \mathbf{l}(C)$. 
Definition

Let $C$ be a convex set in $\mathbb{R}^n$. A hyperplane $H$ is a \textit{supporting hyperplane of $C$} if the following conditions are satisfied:

1. $H$ is closed;
2. $C$ is included in one of the closed half-spaces determined by $H$;
3. $H \cap \mathcal{K}(C) \neq \emptyset$. 

Here, $\mathcal{K}(C)$ denotes the convex hull of $C$. 


Theorem

Let $C$ be a convex set in a linear space $L$. If $\mathbf{l}(C) \neq \emptyset$ and $x_0 \in \partial C$, then there exists a supporting hyperplane $H$ of $C$ such that $x_0 \in H$. 
Theorem

(Separation Theorem) Let $C_1, C_2$ be two non-empty convex sets in $\mathbb{R}^n$ such that $\text{I}(C_1) \neq \emptyset$ and $C_2 \cap \text{I}(C_1) = \emptyset$.

There exists a closed hyperplane $H$, $w'x = a$, separating $C_1$ and $C_2$. In other words, there exists a linear functional $f \in L^*$ such that

$$\sup\{w'x \mid x \in C_1\} \leq \inf\{w'x \mid x \in C_2\},$$

which means that $C_1$ and $C_2$ are located in distinct half-spaces determined by $H$. 

Corollary

Let $C_1, C_2$ be two disjoint subsets of $\mathbb{R}^n$. If $C_1$ is open, then $C_1$ and $C_2$ are separable, that is, that

$$\sup\{w'x \mid x \in C_1\} \leq \inf\{w'x \mid x \in C_2\},$$
Definition

Let $X$ be an open set in $\mathbb{R}^n$ and let $f : X \rightarrow \mathbb{R}$ be a function. The function $f$ is **Gâteaux differentiable** in $x_0$, where $x_0 \in X$ if there exists a linear operator $(D_x f)(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$(D_x f)(x_0)(u) = \lim_{t \rightarrow 0} \frac{f(x_0 + tu) - f(x_0)}{t}$$

for every $u$ such that $x_0 + tu \in X$. The linear operator $(D_x f)(x_0)$ is the **Gâteaux derivative** of $f$ in $x_0$. The **Gâteaux differential** of $f$ at $x_0$ is the linear operator $\delta f(x_0; h)$ given by

$$\delta f(x_0; u) = \lim_{t \rightarrow 0} \frac{f(x_0 + tu) - f(x_0)}{t}.$$
Example

Let \( a \) be a vector in \( \mathbb{R}^n \). Define \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) as \( f(x) = x' a \). We have:

\[
(D_x f)(x_0)(u) = \lim_{t \to 0} \frac{f(x_0 + tu) - f(x_0)}{t} = \lim_{t \to 0} \frac{(x_0 + tu)'a - x_0'a}{t} = \lim_{t \to 0} \frac{tu'a}{t} = u'a.
\]
Example

Let $A \in \mathbb{R}^{n \times n}$ be a matrix and let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be the functional $f(x) = x'Ax$. We have $(Df)(x_0) = x'_0(A + A')$.

By applying the definition of Gâteaux differential we have

$$(Df)(x_0)(u) = \lim_{t \to 0} \frac{f(x_0 + tu) - f(x_0)}{t}$$

$$= \lim_{t \to 0} \frac{(x'_0 + tu')A(x_0 + tu) - x'_0Ax_0}{t}$$

$$= \lim_{t \to 0} \frac{tu'Ax_0 + tx'_0Au + t^2u'Au}{t}$$

$$= u'Ax_0 + x'_0Au = x'_0A'u + x'_0Au$$

$$= x'_0(A + A')u,$$

which yields $(Df)(x_0) = x'_0(A + A')$.

If $A \in \mathbb{R}^{n \times n}$ is symmetric and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is the functional $f(x) = x'Ax$, then $(Df)(x_0) = 2x'_0A$. 
Example

The norm $\| \cdot \| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is not Gâteaux differentiable in $0_n$. Indeed, suppose that $\| \cdot \|$ were differentiable in $0_n$, which would mean that the limit:

$$\lim_{t \to 0} \frac{\| tu \|}{t} = \lim_{t \to 0} \frac{|t|}{t} \| u \|$$

eexists for every $u \in \mathbb{R}^n$, which is contradictory. However, the square of the norm, $\| \cdot \|^2$ is differentiable in $0_n$ because

$$\lim_{t \to 0} \frac{\| tu \|^2}{t} = \lim_{t \to 0} t \| u \| = 0.$$
Example

Consider the norm $\| \cdot \|_1$ on $\mathbb{R}^n$ given by

$$\| x \|_1 = |x_1| + \ldots + |x_n|$$

for $x \in \mathbb{R}^n$. This norm is not Gâteaux differentiable in any point $x_0$ located on an axis. Indeed, let $x_0 = ae_i$ be a point on the $i^{th}$ axis. The limit

$$\lim_{t \to 0} \frac{\| x_0 + tu \|_1 - \| x_0 \|_1}{t} = \lim_{t \to 0} \frac{\| ae_i + tu \|_1 - \| ae_i \|_1}{t}$$

$$= \lim_{t \to 0} \frac{|t||u_1| + \ldots + |t||u_{i-1}| + (|t||u_i| - |a|) + |t||u_{i+1}| + \ldots + |t||u_n|}{t}$$

does not exist, so the norm $\| \cdot \|_1$ is not differentiable in any of these points.
Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and let $h \in \mathbb{R}^n - \{0_n\}$.

The **directional derivative at** $x_0$ **in the direction** $h$ **is** the function $\frac{\partial f}{\partial h}(x_0)$ **given by**

$$\frac{\partial f}{\partial h}(x_0) = \lim_{t \downarrow 0} \frac{f(x_0 + th) - f(x_0)}{t}.$$
$f$ is Gâteaux differentiable at $x_0$ if its directional derivative exists in every direction.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function differentiable at $x_0 \in \mathbb{R}^n$. If \{${e_1, \ldots, e_n}$\} is the standard basis for $\mathbb{R}^n$, then $(Df)(x_0)(e_i)$ is known as the partial derivative of $f$ with respect to $x_i$ and is denoted by $\frac{\partial f}{\partial x_i}(x_0)$. 
Theorem

Let $X$ be an open set in $\mathbb{R}^n$ and let $f : X \rightarrow \mathbb{R}$ be a function. If $f$ is Gâteaux differentiable on $X$, then

$$\| f(u) - f(v) \| \leq \| u - v \| \sup \{ f'(a u + (1 - a)v) \mid a \in [0, 1] \}. $$
Let \( w \in X \) such that \( \| w \| = 1 \) and \( \| f(u) - f(v) \| = (w, f(u) - f(v)) \).
Define the real-valued function \( g \) as \( g(t) = (w, f(u + t(v - u))) \) for \( t \in [0, 1] \).
We have the inequality
\[
\| f(u) - f(v) \| = (w, f(v) - f(u)) = |g(1) - g(0)| \leq \sup\{ |g'(t)| \mid t \in [0, 1] \}.
\]

Since
\[
g'(t) = (w, \text{DER} f(u + t(v - u)) t) \\
= (w, \lim_{r \to 0} \frac{f(u + (t + r)(v - u)) - f(u + t(v - u))}{r}) \\
= (w, f'_{u + t(v-u)}(v - u)) ,
\]
we have \( |g'(t)| \leq \| f'_{u + t(v-u)}(v - u) \| \), hence
\[
|g'(t)| \leq \| f'_{u + t(v-u)}(v - u) \| \leq \| f'_{u + t(v-u)} \| \| v - u \| .
\]
Recall that for $u, v \in \mathbb{R} \cup \{\infty\}$, the sum $u + v$ is always defined. It is useful to extend the notion of convex function by allowing $\infty$ as a value. Thus, if a function $f$ is defined on a subset $S$ of a linear space $L$, $f : S \rightarrow \mathbb{R}$, the \textit{extended-value function} of $f$ is the function $\hat{f} : L \rightarrow \mathbb{R} \cup \{\infty\}$ defined by

$$
\hat{f}(x) = \begin{cases} 
  f(x) & \text{if } x \in S, \\
  \infty & \text{otherwise},
\end{cases}
$$

If a function $f : S \rightarrow \mathbb{R}$ is convex, where $S \subseteq L$ is a convex set, then its extended-value function $\hat{f}$ satisfies the inequality that defines convexity $\hat{f}((1 - t)x + ty) \leq (1 - t)\hat{f}(x) + t\hat{f}(y)$ for every $x, y \in L$ and $t \in [0, 1]$, if we adopt the convention that $0 \cdot \infty = 0$. 
Definition

The trivial convex function is the function $f_\infty : S \rightarrow \mathbb{R} \cup \{\infty\}$ defined by $f(x) = \infty$ for every $x \in S$.

A extended-value convex function $\hat{f} : S \rightarrow \mathbb{R} \cup \{\infty\}$ is properly convex or a proper function if $\hat{f} \neq f_\infty$.

The domain of a function $f : S \rightarrow \mathbb{R} \cup \{\infty\}$ is the set $\text{Dom}(f) = \{x \in S \mid f(x) < \infty\}$.
Example

Let $f : (0, \infty) \rightarrow \mathbb{R}$ be defined by $f(x) = x^2$. The definition domain of $f$ is clearly convex and we have:

\[
f((1 - t)x_1 + tx_2) = ((1 - t)x_1 + tx_2)^2
= (1 - t)^2 x_1^2 + t^2 x_2^2 + 2(1 - t)tx_1x_2.
\]

Therefore,

\[
f((1 - t)x_1 + tx_2) - (1 - t)f(x_1) - tf(x_2)
= (1 - t)^2 x_1^2 + t^2 x_2^2 + 2(1 - t)tx_1x_2 - (1 - t)x_1^2 - tx_2^2
= -t(1 - t)(x_1 - x_2)^2 \leq 0,
\]

which implies that $f$ is indeed convex.
Example

The function \( f : \mathbb{R} \rightarrow \mathbb{R} \) defined by \( f(x) = |a - xb| \) is convex because

\[
\begin{align*}
    f((1 - t)x_1 + tx_2) &= |a - ((1 - t)x_1 + tx_2)b| \\
    &= |a(1 - t) + at - ((1 - t)x_1 + tx_2)b| \\
    &= |(1 - t)(a - x_1 b) + t(a - x_2 b)| \\
    &\leq |(1 - t)(a - x_1 b)| + |t(a - x_2 b)| = (1 - t)f(x_1) + tf(x_2)
\end{align*}
\]

for \( t \in [0, 1] \).
Example

Any norm $\nu$ on a real linear space $L$ is convex. Indeed, for $t \in [0, 1]$ we have

$$\nu(tx + (1 - t)y) \leq \nu(tx) + \nu((1 - t)y) = t\nu(x) + (1 - t)\nu(y)$$

for $x, y \in L$. 
It is easy to verify that any linear combination of convex functions with non-negative coefficients defined on a real linear space $L$ (of functions convex at $x_0 \in L$) is a convex function (a function convex at $x_0$).
Example

Let $A \in \mathbb{R}^{n \times n}$ be a matrix. If $A$ is a positive matrix then the function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by $f(x) = x'Ax$ for $x \in \mathbb{R}^n$ is convex on $\mathbb{R}^n$. Let $t \in [0, 1]$ and let $x, y \in \mathbb{R}^n$. By hypothesis we have

$$(t - t^2)(x - y)'A(x - y) \geq 0$$

for $x, y \in \mathbb{R}^n$ because $t - t^2 \geq 0$. Therefore,

$$(1 - t)x'Ax + ty'Ay$$

$$= x'Ax + tx'(y - x) + t(y - x)'Ax + t(y - x)'A(y - x)$$

$$\geq x'Ax + tx'(y - x) + t(y - x)'Ax + t^2(y - x)'A(y - x)$$

$$= (x + t(y - x))'A(x + t(y - x))$$

for $t \in [0, 1]$, which proves the convexity of $f$. 
Theorem

Let \((a, b)\) be an open interval of \(\mathbb{R}\) and let \(f : (a, b) \rightarrow \mathbb{R}\) be a differentiable function on \((a, b)\). Then, \(f\) is convex on \((a, b)\) if and only if \(f(y) \geq f(x) + f'(x)(y - x)\) for every \(x, y \in (a, b)\).
Proof

Suppose that $f$ is convex on $(a, b)$. Then, for $x, y \in (a, b)$ we have

$$f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y)$$

for $t \in [0, 1]$. Therefore, for $t < 1$ we have

$$f(y) \geq f(x) + \frac{f(x + t(y - x)) - f(x)}{t(y - x)}(y - x).$$

When $t \to 0$ we obtain $f(y) \geq f(x) + f'(x)(y - x)$.

Conversely, suppose that $f(y) \geq f(x) + f'(x)(y - x)$ for every $x, y \in (a, b)$ and let $z = (1 - t)x + ty$. We have

$$f(x) \geq f(z) + f'(z)(x - z),$$

$$f(y) \geq f(z) + f'(z)(y - z).$$

By multiplying the first inequality by $1 - t$ and the second by $t$ we obtain

$$(1 - t)f(x) + tf(y) \geq f(z),$$

which shows that $f$ is convex.
Theorem

Let $S$ be a convex subset of $\mathbb{R}^n$ and let $f : S \rightarrow \mathbb{R}$ be a Gâteaux differentiable function on $S$. Then, $f$ is convex on $S$ if and only if $f(y) \geq f(x) + (\nabla f)(x)'(y - x)$ for every $x, y \in S$. 
Corollary

Let $S$ be a convex subset of $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$ be a Gâteaux differentiable function on $S$. If $(\nabla f)(x_0)'(x - x_0) \geq 0$ for every $x \in S$, then $f(x_0)$ is a minimum for $f$ in $S$. 
Example

Let $S = K_{\text{conv}}\{a_1, \ldots, a_m\} \subseteq \mathbb{R}^n$ and let $f : S \rightarrow \mathbb{R}$ be the linear function defined by $f(x) = c'x$. We have $(\nabla f)(x) = c$.

An application of the previous Corollary shows that if $x_0$ minimizes $f$ on $S$ then $c'(a_i - x_0) \geq 0$ for every $1 \leq i \leq n$.

Since $x \in S$ if and only if $x = \sum_{i=1}^{n} b_i a_i$, where $\sum_{i=1}^{m} b_i = 1$, it follows that we have:

$$c' \left( \sum_{i=1}^{n} b_i a_i - x_0 \right) \geq 0.$$  

This condition is satisfied with

$$b_j = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{otherwise} \end{cases}$$

when $x_0 = a_i$. Thus, each of the extreme points $a_i$ of $S$ achieves a minimum for $f$ on $S$. 


Theorem

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex, differentiable function. Any critical point $x_0$ of $f$ is a global minimum for $f$. 
Proof

Let $x_0$ be a critical point for $f$. Suppose that $x_0$ is not a global minimum for $f$. Then, there exists $z$ such that $f(z) < f(x_0)$. Since $f$ is differentiable in $x_0$, we have

$$(\nabla f)'_{x_0}(z - x_0) = \frac{d}{dt} f(x_0 + t(z - x_0))_{t=0}$$

$$= \lim_{t \to 0} \frac{f(x_0 + t(z - x_0)) - f(x_0)}{t}$$

$$= \lim_{t \to 0} \frac{f(tz + (1 - t)x_0) - f(x_0)}{t}$$

$$\leq \frac{tf(z) + (1 - t)f(x_0) - f(x_0)}{t}$$

$$= \frac{t(f(z) - tf(x_0))}{t} < 0,$$

which implies $\nabla f|_{x_0} \neq 0_n$, thus contradicting the fact that $x_0$ is a critical point.