Convex Functions on \( \mathbb{R} \)

If \( I = [a, b] \) and \( f : [a, b] \to \mathbb{R} \), let

\[
s(I) = \frac{f(b) - f(a)}{b - a}.
\]

If \( I_1 = [a_1, b_1] \) and \( I_2 = [a_2, b_2] \) are intervals, we say \( I_1 \leq I_2 \) if \( a_1 \leq a_2 \) and \( b_1 \leq b_2 \).

**Theorem.** Let \( f : (a, b) \to \mathbb{R} \) be a function. Then \( f \) is convex if and only if \( s(I) \leq s(J) \) for all subintervals \( I, J \) satisfying \( I \leq J \).

(done in class.)

A function is *Lipschitz* on \([a, b]\) if \(|f(x) - f(x')| \leq C|x - x'| \) for \( a \leq x, x' \leq b \). Clearly a Lipschitz function is continuous.

**Corollary 1.** If \( f \) is convex on an open interval \( I \) containing \([a, b]\), then \( f \) is Lipschitz on \([a, b]\).

**Proof.** Pick \( a', b' \) in \( I \) satisfying \( a' < a < b < b' \). Then for \( x, x' \in [a, b] \), we have \([a', a] \leq [x, x'] \leq [b, b'] \) hence

\[
s([a', a]) \leq \frac{f(x') - f(x)}{x' - x} \leq s([b, b']).
\]

Thus \( f \) is Lipschitz on \([a, b]\).

**Corollary 2.** If \( f : (a, b) \to \mathbb{R} \) is convex, then

\[
f'_{\pm}(x) = \lim_{h \to 0^{\pm}} \frac{f(x + h) - f(x)}{h}
\]

exists and \( f'_-(x) \leq f'_+(x) \), for all \( a < x < b \).

**Proof.** For \( h \) real, let \( I_h \) be the interval with endpoints \( x \) and \( x + h \). By the Theorem, \( s(I_h) \) is an increasing function of \( h \), so \( f'_\pm(x) \) both exist. Now let \( I = [x - h, x] \) and \( J = [x, x + k] \). Then \( I \leq J \) so by the Theorem

\[
\frac{f(x) - f(x - h)}{h} = s(I) \leq s(J) = \frac{f(x + k) - f(x)}{k}.
\]
Taking limits as \( h, k \to 0^+ \), we get the result.

**Corollary 3.** If \( f : (a, b) \to \mathbb{R} \) is convex and \( c \in (a, b) \), then

\[
f(x) \geq f(c) + p(x - c), \quad a < x < b,
\]

iff \( f'_-(c) \leq p \leq f'_+(c) \).

Note: A line \( y = f(c) + p(x - c) \) satisfying the above inequality is a support line at \( c \).

**Proof.** Let \( I_x \) denote the interval joining \( c \) and \( x \). Then, by the Theorem, \( s(I_{c+h}) \leq s(I_x) \) for \( c < c + h < x \). Let \( h \to 0^+ \) to get \( f'_+(c) \leq s(I_x) \) for \( x > c \). If \( p \leq f'_+(c) \), this gives

\[
f(x) \geq f(c) + p(x - c), \quad x > c.
\]

Similarly, we have \( s(I_{c-h}) \geq s(I_x) \) for \( x < c - h < c \). Let \( h \to 0^+ \) to get \( f'_-(c) \geq s(I_x) \) for \( x < c \). If \( p \geq f'_-(c) \), this gives

\[
f(x) \geq f(c) + p(x - c), \quad x < c.
\]

Thus if \( p \in [f'_-(c), f'_+(c)] \), the inequality holds on \( (a, b) \). Conversely, if the inequality holds for a given \( p \) and for all \( x \) in \( (a, b) \), dividing by \( x - c \) for \( x > c \) and taking the limit \( x \to c^+ \), we get \( f'_+(c) \geq p \). Similarly, dividing by \( x - c \) for \( x < c \) and taking the limit \( x \to c^- \), we get \( f'_-(c) \leq p \).

**Corollary 4.** If \( f : (a, b) \to \mathbb{R} \) is convex and \( a < x < x' < b \), then

\[
f'_\pm(x) \leq f'_\pm(x').
\]

**Proof.** Let \( I \) and \( I' \) denote the intervals joining \( x \) to \( x + h \) and \( x' \) to \( x' + h' \), for \( h, h' \) real. Then for \( h, h' \) small, we have \( I \leq I' \) so \( s(I) \leq s(I') \). Taking limits yields the result.
Corollary 5. If $f : (a, b) \to \mathbb{R}$ is convex and differentiable, then $f'$ is continuous on $(a, b)$.

Proof. If $[x - \delta, x + \delta] \subset (a, b)$, then for $0 < h < \delta$, by the previous Corollaries, $f'(x) \leq f'(x + h) \leq s([x + h, x + \delta])$. Let $h \to 0^+$ then $\delta \to 0^+$, we get $f'$ is continuous from the right. Similarly from the left.

Corollary 6. If $f$ is convex on an open interval containing $[a, b]$, then

$$f(b) - f(a) = \int_a^b f'_\pm(x) \, dx.$$  

Proof. If $a = x_0 < x_1 < \cdots < x_n = b$ is a partition, then

$$\int_a^b f'_+(x) \, dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f'_+(x) \, dx$$

$$\geq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f'_+(x_{i-1}) \, dx$$

$$= \sum_{i=2}^n f'_+(x_{i-1})(x_i - x_{i-1}) + f'_+(x_0)(x_1 - x_0)$$

$$\geq \sum_{i=2}^n (f(x_{i-1}) - f(x_{i-2})) + f'_+(x_0)(x_1 - x_0)$$

$$= f(x_{n-1}) - f(x_0) + f'_+(x_0)(x_1 - x_0).$$

Taking the limit as $\Delta x_i \to 0$, we get $\int_a^b f'_+(x) \, dx \geq f(b) - f(a)$. Similarly from the other side.
Convex Functions on $\mathbb{R}^n$

Throughout $G$ denotes a convex open set in $\mathbb{R}^n$.

**Theorem.** A convex function $f : G \to \mathbb{R}$ is Lipschitz on any compact convex subset $K \subset \subset G$.

Let $Q$ be a compact cube in $G$ and let $\partial_i^\pm f(x)$ denote the partials at $x$. Let $Q^0, Q^1, Q^2, \ldots$, denote the vertices, edges, faces, \ldots, of $Q$. Since $Q^0$ is a finite set, these partials are bounded on $Q^0$ by some $C$. From the one-dimensional case, these partials are bounded on $Q^1$ by the same $C$. By the one-dimensional case again, these partials are bounded on $Q^2$ by the same $C$. Proceeding in this way, we see that these partials are all bounded on $Q$. In particular, these partials are bounded on the boundary $\partial Q$ by some $C$.

Now if $x, x' \in Q$ and $x' - x$ is parallel to a coordinate axis, by the one-dimensional case,

$$\frac{f(x) - f(x')}{|x - x'|}$$

is bounded by the partials on $\partial Q$, hence bounded by $C$. In general, any $x$ and $x'$ in $Q$ can be joined by a polygonal path with $n$ segments parallel to the axes, hence the above ratio is bounded on $Q$ by $nC$. This shows $f$ is Lipschitz on any compact cube $Q$. If $K$ is a compact convex subset of $G$, cover $K$ by finitely many cubes $Q_i$. Then for $x, x' \in K$, join $x, x'$ by a line segment and apply the previous step to get $f$ is Lipschitz on $K$.  

A vector $p \in \mathbb{R}^n$ is a support plane at $a \in G$ if

$$f(x) \geq f(a) + p \cdot (x - a), \quad x \in G.$$

Here $p \cdot v$ means inner product.

**Theorem.** Fix $a \in G$. Then there exists a support plane at $a$. 
For \( v \in \mathbb{R}^n \), by the one-dimensional case, \((f(a + tv) - f(a))/t \) is an increasing function of \( t > 0 \), hence

\[
N(v) \equiv \lim_{{t \to 0^+}} \frac{f(a + tv) - f(a)}{t} = \inf_{{t > 0}} \frac{f(a + tv) - f(a)}{t}
\]

exists. Using convexity one checks \( N \) satisfies

1. \( N(sv) = sN(v) \) for \( s > 0 \), \( (N \) is positively homogeneous) 
2. \( N(v + w) \leq N(v) + N(w) \). \( (N \) is subadditive) 

(Note \( N \) is almost a norm.) If we find a \( p \in \mathbb{R}^n \) satisfying

\[(*) \quad p \cdot v \leq N(v), \quad v \in \mathbb{R}^n, \]

we will be done, since this implies

\[
p \cdot v \leq \frac{f(a + tv) - f(a)}{t}, \quad t > 0
\]

which implies the result by choosing \( t = 1 \) and \( x = a + v \). (*) follows from the Hahn-Banach theorem below (choose \( V_0 = \{0\} \)). \( \square \)

Let \( f : G \to \mathbb{R} \) be convex. The subdifferential \( \partial f(a) \) of \( f \) at \( a \) is the set of support planes \( p \) at \( a \).

**Theorem.** Let \( f : G \to \mathbb{R} \) be convex and let \( a \in G \). Then \( \partial f(a) \) is a nonempty compact convex set in \( \mathbb{R}^n \).

We showed already that \( \partial f(a) \) is nonempty. The rest is a straightforward consequence of convexity.

**Hahn-Banach Theorem (in \( \mathbb{R}^n \)).** Suppose we are given \( N : \mathbb{R}^n \to \mathbb{R} \) positively homogeneous and subadditive and \( p_0 \in \mathbb{R}^n \) and a vector subspace \( V_0 \subset \mathbb{R}^n \) such that \( p_0 \cdot v \leq N(v) \) for \( v \in V_0 \). Then there is a \( p \in \mathbb{R}^n \) such that \( p \cdot v \leq N(v) \) for \( v \in \mathbb{R}^n \) and \( p \cdot v = p_0 \cdot v \) for \( v \in V_0 \).

To see this, choose unit vectors \( e_1, e_2, \ldots, e_r \) orthogonal to each other and to \( V_0 \) such that the vector subspaces \( V_i = V_0 + \mathbb{R}e_1 + \cdots + \mathbb{R}e_i \) are
distinct and $V_0 \subset V_1 \subset \cdots \subset V_r = \mathbb{R}^n$. It is enough to find $p_1$ such that $p_1 \cdot v_1 \leq N(v)$ for $v_1 \in V_1$ and $p_1 \cdot v_0 = p_0 \cdot v_0$ for $v_0 \in V_0$, for then we can proceed by induction.

Without loss of generality, we may assume $p_0 \in V_0$; otherwise project $p_0$ onto $V_0$. If $v_0, v'_0 \in V_0$, then

$$p_0 \cdot v_0 + p_0 \cdot v'_0 = p_0 \cdot (v_0 + v'_0) \leq N(v_0 + v'_0) \leq N(v_0 + e_1) + N(v'_0 - e_1).$$

Thus

$$p_0 \cdot v - N(v_0 + e_1) \leq -p_0 \cdot v'_0 + N(v'_0 - e_1).$$

Since $v_0, v'_0$ were arbitrary in $V_0$, we get

$$\sup_{v_0 \in V_0} (p_0 \cdot v_0 - N(v_0 + e_1)) \leq \inf_{v_0 \in V_0} (-p_0 \cdot v_0 + N(v_0 - e_1)).$$

Thus there exists a $c \in \mathbb{R}$ satisfying

\[
\text{(**) } \sup_{v_0 \in V_0} (p_0 \cdot v_0 - N(v_0 + e_1)) \leq c \leq \inf_{v_0 \in V_0} (-p_0 \cdot v_0 + N(v_0 - e_1)).
\]

Define $p_1 = p_0 - ce_1$. Clearly $p_1 \cdot v_0 = p_0 \cdot v_0$ for $v_0 \in V_0$. Now, for $v_0 \in V_0$, apply the left side of (**) with $v_0/t$ replacing $v_0$ and $t > 0$, to get

$$p_1 \cdot (v_0 + te_1) = p_0 \cdot v_0 - tc = t \left( p_0 \cdot \left( \frac{v_0}{t} \right) - c \right)$$

$$\leq t \left( N \left( \frac{v_0}{t} + e_1 \right) + c - c \right) = N(v_0 + te_1).$$

Similarly, apply the right side of (**) with $v_0/t$ replacing $v_0$ and $t > 0$, to get

$$p_1 \cdot (v_0 - te_1) = p_0 \cdot v_0 + tc = t \left( p_0 \cdot \left( \frac{v_0}{t} \right) + c \right)$$

$$\leq t \left( N \left( \frac{v_0}{t} - e_1 \right) - c + c \right) = N(v_0 - te_1).$$

Since every vector in $V_1$ is of the form $v_0 + tv_1$ for $t \in \mathbb{R}$, we get $p_1 \cdot v \leq N(v)$ for all $v \in V_1$. □

Note: The Hahn-Banach theorem is true on any normed vector space (not just $\mathbb{R}^n$) and is widely useful.