Math 408A
Convexity

Convexity

Definition: *A set $C \subset \mathbb{R}^n$ is said to be convex if for every $x, y \in C$ and $\lambda \in [0, 1]$ one has*

$$(1 - \lambda)x + \lambda y \in C \ .$$

Convexity

Definition: *A set $C \subset \mathbb{R}^n$ is said to be convex if for every $x, y \in C$ and $\lambda \in [0,1]$ one has*

$$(1-\lambda)x + \lambda y \in C\ .$$

Definition: *A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be convex if for every two points $x_1, x_2 \in \mathbb{R}^n$ and $\lambda \in [0,1]$ we have*

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2).$$

Convexity

Definition: *A set* $C \subset \mathbb{R}^n$ *is said to be convex if for every* $x, y \in C$ *and* $\lambda \in [0,1]$ *one has*

$$(1-\lambda)x + \lambda y \in C \ .$$

Definition: *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is said to be convex if for every two points* $x_1, x_2 \in \mathbb{R}^n$ *and* $\lambda \in [0,1]$ *we have*

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2).$$

The function f *is said to be strictly convex if for every two distinct points* $x_1, x_2 \in \mathbb{R}^n$ *and* $\lambda \in [0,1]$ *we have*

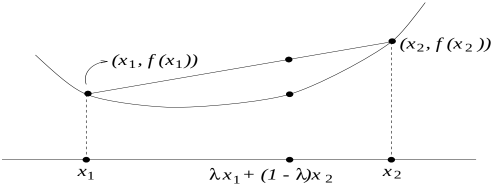
$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda)f(x_2).$$

Convexity

The secant line connecting $(x_1, f(x_1))$ and $(x_2, f(x_2))$ lies above the graph of f.

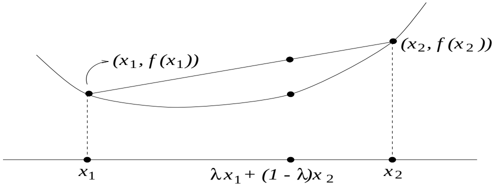
Convexity

The secant line connecting $(x_1, f(x_1))$ and $(x_2, f(x_2))$ lies above the graph of f.



Convexity

The secant line connecting $(x_1, f(x_1))$ and $(x_2, f(x_2))$ lies above the graph of f.



That is, the set

$$\mathrm{epi}\,(f) = \{(x, \mu) \,:\, f(x) \le \mu\},$$

called the *epi-graph* of f, is a convex set.

Convexity

Definition: *A function* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} = \bar{\mathbb{R}}$ *is said to be convex if the set* $epi\,(f) = \{(x, \mu)\ :\ f(x) \leq \mu\}$ *is a convex set. We also define the essential domain of f to be the set*

$$dom\,(f) = \{x\ :\ f(x) < +\infty\}\ .$$

Convexity

Definition: *A function* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\} = \bar{\mathbb{R}}$ *is said to be convex if the set* $epi(f) = \{(x, \mu) \,:\, f(x) \le \mu\}$ *is a convex set. We also define the essential domain of* f *to be the set*

$$dom(f) = \{x \,:\, f(x) < +\infty\} \,.$$

Examples of convex functions are

$$c^T x, \quad \|x\|, \quad x^T H x \quad \text{when } H \text{ is psd.}$$

Convexity and Optimality

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex. If $\overline{x} \in \mathbb{R}^n$ is a local minimum for f, then $\overline{x}$ is a global minimum for f. If, in addition, f is strictly convex, then $\bar{x}$ is the unique global minimizer.*

Convexity and Optimality

Proof: Suppose $f(\widehat{x}) < f(\overline{x})$. Let $\epsilon > 0$ be such that

$$f(\overline{x}) \leq f(x) \quad \text{whenever} \quad \|x - \overline{x}\| \leq \epsilon \quad \text{and} \quad \epsilon < 2\|\overline{x} - \widehat{x}\| \ .$$

Convexity and Optimality

Proof: Suppose $f(\widehat{x}) < f(\overline{x})$. Let $\epsilon > 0$ be such that

$$f(\overline{x}) \leq f(x) \quad \text{whenever} \quad \|x - \overline{x}\| \leq \epsilon \quad \text{and} \quad \epsilon < 2\|\overline{x} - \widehat{x}\| \ .$$

Set $\lambda := \epsilon(2\|\overline{x} - \widehat{x}\|)^{-1} < 1$ and $x_\lambda := \overline{x} + \lambda(\widehat{x} - \overline{x})$.

Convexity and Optimality

Proof: Suppose $f(\widehat{x}) < f(\overline{x})$. Let $\epsilon > 0$ be such that

$$f(\overline{x}) \leq f(x) \quad \text{whenever} \quad \|x - \overline{x}\| \leq \epsilon \quad \text{and} \quad \epsilon < 2\|\overline{x} - \widehat{x}\| \ .$$

Set $\lambda := \epsilon(2\|\overline{x} - \widehat{x}\|)^{-1} < 1$ and $x_\lambda := \overline{x} + \lambda(\widehat{x} - \overline{x})$.
Then $\|x_\lambda - \overline{x}\| \leq \epsilon/2$ and

$$f(x_\lambda) \leq (1 - \lambda)f(\overline{x}) + \lambda f(\widehat{x}) < f(\overline{x}).$$

Convexity and Optimality

Proof: Suppose $f(\widehat{x}) < f(\overline{x})$. Let $\epsilon > 0$ be such that

$$f(\overline{x}) \le f(x) \quad \text{whenever} \quad \|x - \overline{x}\| \le \epsilon \quad \text{and} \quad \epsilon < 2\|\overline{x} - \widehat{x}\| \ .$$

Set $\lambda := \epsilon(2\|\overline{x} - \widehat{x}\|)^{-1} < 1$ and $x_\lambda := \overline{x} + \lambda(\widehat{x} - \overline{x})$.
Then $\|x_\lambda - \overline{x}\| \le \epsilon/2$ and

$$f(x_\lambda) \le (1-\lambda)f(\overline{x}) + \lambda f(\widehat{x}) < f(\overline{x}).$$

This contradicts the choice of ϵ and so no such $\widehat{x}$ exists.

Directional Derivatives and Convexity

Lemma: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex (not necessarilty differentiable).*

Directional Derivatives and Convexity

Lemma: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex (not necessarilty differentiable).*

1. *Given $x, d \in \mathbb{R}^n$ the difference quotient*

$$\frac{f(x + td) - f(x)}{t}$$

is a non-decreasing function of t on $(0, +\infty)$.

Directional Derivatives and Convexity

Lemma: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex (not necessarilty differentiable).*

1. *Given $x, d \in \mathbb{R}^n$ the difference quotient*

$$\frac{f(x + td) - f(x)}{t}$$

   is a non-decreasing function of t on $(0, +\infty)$.

2. *For every $x, d \in \mathbb{R}^n$ the directional derivative $f'(x; d)$ always exists and is given by*

$$f'(x; d) := \inf_{t>0} \frac{f(x + td) - f(x)}{t}.$$

Directional Derivatives and Convexity

We first assume (1) is true and show (2).

Directional Derivatives and Convexity

We first assume (1) is true and show (2).
Recall that

$$f'(x;d) := \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t}.$$

Now if the difference quotient on the right is non-decreasing in t on $(0, +\infty)$, then the limit is necessarily given by the infimum of these ratios.

Directional Derivatives and Convexity

We first assume (1) is true and show (2).
Recall that

$$f'(x; d) := \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}.$$

Now if the difference quotient on the right is non-decreasing in t on $(0, +\infty)$, then the limit is necessarily given by the infimum of these ratios.
This infimum always exists and so $f'(x; d)$ always exists and is given by the infimum.

Directional Derivatives and Convexity

We now show that $\frac{f(x+td)-f(x)}{t}$ is non-decreasing for $t > 0$.

Directional Derivatives and Convexity

We now show that $\frac{f(x+td)-f(x)}{t}$ is non-decreasing for $t > 0$.
Let $x, d \in \mathbb{R}^n$ and let $0 < t_1 < t_2$. Then

Directional Derivatives and Convexity

We now show that $\frac{f(x+td)-f(x)}{t}$ is non-decreasing for $t > 0$.
Let $x, d \in \mathbb{R}^n$ and let $0 < t_1 < t_2$. Then

$$f(x + t_1 d) \quad = \quad f\left(x + \left(\frac{t_1}{t_2}\right) t_2 d\right)$$

Directional Derivatives and Convexity

We now show that $\frac{f(x+td)-f(x)}{t}$ is non-decreasing for $t>0$.
Let $x, d \in \mathbb{R}^n$ and let $0 < t_1 < t_2$. Then

$$\begin{aligned} f(x+t_1 d) &= f\left(x + \left(\frac{t_1}{t_2}\right) t_2 d\right) \\ &= f\left[\left(1-\left(\frac{t_1}{t_2}\right)\right)x + \left(\frac{t_1}{t_2}\right)(x+t_2 d)\right] \end{aligned}$$

Directional Derivatives and Convexity

We now show that $\frac{f(x+td)-f(x)}{t}$ is non-decreasing for $t>0$.
Let $x, d \in \mathbb{R}^n$ and let $0 < t_1 < t_2$. Then

$$\begin{aligned} f(x+t_1 d) &= f\left(x + \left(\frac{t_1}{t_2}\right) t_2 d\right) \\ &= f\left[\left(1 - \left(\frac{t_1}{t_2}\right)\right) x + \left(\frac{t_1}{t_2}\right)(x + t_2 d)\right] \\ &\leq \left(1 - \frac{t_1}{t_2}\right) f(x) + \left(\frac{t_1}{t_2}\right) f(x + t_2 d). \end{aligned}$$

Directional Derivatives and Convexity

We now show that $\frac{f(x+td)-f(x)}{t}$ is non-decreasing for $t > 0$.
Let $x, d \in \mathbb{R}^n$ and let $0 < t_1 < t_2$. Then

$$\begin{aligned} f(x+t_1 d) &= f\left(x + \left(\frac{t_1}{t_2}\right) t_2 d\right) \\ &= f\left[\left(1 - \left(\frac{t_1}{t_2}\right)\right) x + \left(\frac{t_1}{t_2}\right)(x + t_2 d)\right] \\ &\le \left(1 - \frac{t_1}{t_2}\right) f(x) + \left(\frac{t_1}{t_2}\right) f(x + t_2 d). \end{aligned}$$

Hence

$$\frac{f(x+t_1 d) - f(x)}{t_1} \le \frac{f(x+t_2 d) - f(x)}{t_2}.$$

The Subdifferential Inequality

If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then for all $x \in \mathrm{dom}\,(f)$ and $y \in \mathbb{R}^n$,

$$f'(x; (y - x)) = \inf_{t>0} \frac{f(x + t(y - x)) - f(x)}{t} \; .$$

The Subdifferential Inequality

If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then for all $x \in \mathrm{dom}\,(f)$ and $y \in \mathbb{R}^n$,

$$f'(x;(y-x)) = \inf_{t>0} \frac{f(x+t(y-x)) - f(x)}{t} .$$

Plug in $t = 1$ on the right hand side to get

$$f(y) \geq f(x) + f'(x;(y-x)) \qquad \forall\, y \in \mathbb{R}^n .$$

The Subdifferential Inequality

If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then for all $x \in \operatorname{dom}(f)$ and $y \in \mathbb{R}^n$,

$$f'(x;(y-x)) = \inf_{t>0} \frac{f(x+t(y-x)) - f(x)}{t} \; .$$

Plug in $t = 1$ on the right hand side to get

$$f(y) \geq f(x) + f'(x;(y-x)) \qquad \forall \; y \in \mathbb{R}^n \; .$$

We will call this inequality the *subdifferential inequality*.

The Subdifferential Inequality

If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then for all $x \in \mathrm{dom}\,(f)$ and $y \in \mathbb{R}^n$,

$$f'(x;(y-x)) = \inf_{t>0} \frac{f(x+t(y-x)) - f(x)}{t} \; .$$

Plug in $t = 1$ on the right hand side to get

$$f(y) \geq f(x) + f'(x;(y-x)) \qquad \forall\; y \in \mathbb{R}^n \; .$$

When f is differentiable at x, this inequality becomes

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \qquad \forall\; y \in \mathbb{R}^n \; .$$

The Subdifferential Inequality

If $f : \mathbb{R}^n \to \mathbb{R}$ is convex, then for all $x \in \mathrm{dom}\,(f)$ and $y \in \mathbb{R}^n$,

$$f'(x;(y-x)) = \inf_{t>0} \frac{f(x+t(y-x)) - f(x)}{t} \; .$$

Plug in $t = 1$ on the right hand side to get

$$f(y) \geq f(x) + f'(x;(y-x)) \qquad \forall\, y \in \mathbb{R}^n \; .$$

When f is differentiable at x, this inequality becomes

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \qquad \forall\, y \in \mathbb{R}^n \; .$$

This inequality states that the tangent plane to the graph of f lies entirely below the graph of f.

Convexity and Optimality Conditions

Theorem: *Let* $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ *be convex (not necessarilty differentiable) and let* $\bar{x} \in \mathrm{dom}\,(f)$*. Then the following three statements are equivalent.*

Convexity and Optimality Conditions

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex (not necessarilty differentiable) and let $\bar{x} \in \mathrm{dom}\,(f)$. Then the following three statements are equivalent.*

(i) *$\bar{x}$ is a local solution to $\min_{x \in \mathbb{R}^n} f(x)$.*

Convexity and Optimality Conditions

Theorem: *Let* $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ *be convex (not necessarilty differentiable) and let* $\bar{x} \in \operatorname{dom}(f)$*. Then the following three statements are equivalent.*

(i) $\bar{x}$ *is a local solution to* $\min_{x \in \mathbb{R}^n} f(x)$*.*

(ii) $f'(\bar{x}; d) \geq 0$ *for all* $d \in \mathbb{R}^n$*.*

Convexity and Optimality Conditions

Theorem: *Let* $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ *be convex (not necessarilty differentiable) and let* $\bar{x} \in \mathrm{dom}\,(f)$*. Then the following three statements are equivalent.*

(i) $\bar{x}$ *is a local solution to* $\min_{x\in\mathbb{R}^n} f(x)$*.*

(ii) $f'(\bar{x}; d) \geq 0$ *for all* $d \in \mathbb{R}^n$*.*

(iii) $\bar{x}$ *is a global solution to* $\min_{x\in\mathbb{R}^n} f(x)$*.*

Convexity and Optimality Conditions

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex (not necessarilty differentiable) and let $\bar{x} \in \mathrm{dom}\,(f)$. Then the following three statements are equivalent.*

(i) *$\bar{x}$ is a local solution to $\min_{x\in\mathbb{R}^n} f(x)$.*

(ii) *$f'(\bar{x}; d) \geq 0$ for all $d \in \mathbb{R}^n$.*

(iii) *$\bar{x}$ is a global solution to $\min_{x\in\mathbb{R}^n} f(x)$.*

Proof: (i)⇒(ii): Done.

Convexity and Optimality Conditions

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex (not necessarilty differentiable) and let $\bar{x} \in \mathrm{dom}\,(f)$. Then the following three statements are equivalent.*

(i) *$\bar{x}$ is a local solution to $\min_{x\in\mathbb{R}^n} f(x)$.*

(ii) *$f'(\bar{x}; d) \geq 0$ for all $d \in \mathbb{R}^n$.*

(iii) *$\bar{x}$ is a global solution to $\min_{x\in\mathbb{R}^n} f(x)$.*

Proof: (i)⇒(ii): Done.
(ii)⇒(iii): Follows from the subdifferential inequality;

$$f(y) \geq f(\bar{x}) + f'(\bar{x}; (y-x)) \geq f(\bar{x}) \qquad \forall\, y \in \mathbb{R}^n\,.$$

Convexity and Optimality Conditions

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex (not necessarilty differentiable) and let $\bar{x} \in \mathrm{dom}\,(f)$. Then the following three statements are equivalent.*

(i) *$\bar{x}$ is a local solution to $\min_{x \in \mathbb{R}^n} f(x)$.*

(ii) *$f'(\bar{x}; d) \geq 0$ for all $d \in \mathbb{R}^n$.*

(iii) *$\bar{x}$ is a global solution to $\min_{x \in \mathbb{R}^n} f(x)$.*

Proof: (i)⇒(ii): Done.
(ii)⇒(iii): Follows from the subdifferential inequality;

$$f(y) \geq f(\bar{x}) + f'(\bar{x}; (y - x)) \geq f(\bar{x}) \qquad \forall\ y \in \mathbb{R}^n\ .$$

(iii)⇒(i): Obvious.

Convexity and Optimality Conditions

Theorem: *Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and suppose that $\bar{x} \in \mathbb{R}^n$ is a point at which f is differentiable. Then $\bar{x}$ is a global minimum of f if and only if $\nabla f(\bar{x}) = 0$.*

Checking Convexity

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$.*

If f is differentiable on $\mathbb{R}^n$, then the following statements are equivalent.

Checking Convexity

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$.*

If f is differentiable on $\mathbb{R}^n$, then the following statements are equivalent.

(1) *f is convex,*

Checking Convexity

Theorem: *Let* $f : \mathbb{R}^n \to \bar{\mathbb{R}}$.

If f *is differentiable on* $\mathbb{R}^n$, *then the following statements are equivalent.*

(1) f *is convex,*
(2) $f(y) \geq f(x) + \nabla f(x)^T (y - x)$ *for all* $x, y \in \mathbb{R}^n$

Checking Convexity

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$.*

If f is differentiable on $\mathbb{R}^n$, then the following statements are equivalent.

(1) *f is convex,*
(2) $f(y) \geq f(x) + \nabla f(x)^T(y - x)$ *for all* $x, y \in \mathbb{R}^n$
(3) $(\nabla f(x) - \nabla f(y))^T(x - y) \geq 0$ *for all* $x, y \in \mathbb{R}^n$.

Checking Convexity

Theorem: *Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$.*

If f is differentiable on $\mathbb{R}^n$, then the following statements are equivalent.

(1) *f is convex,*
(2) *$f(y) \geq f(x) + \nabla f(x)^T(y - x)$ for all $x, y \in \mathbb{R}^n$*
(3) *$(\nabla f(x) - \nabla f(y))^T(x - y) \geq 0$ for all $x, y \in \mathbb{R}^n$.*

If f is twice differentiable then f is convex if and only if $\nabla^2 f(x)$ is positive semi-definite for all $x \in \mathbb{R}^n$.

$f(y) \geq f(x) + \nabla f(x)^T(y-x) \ \forall x, y \in \mathbb{R}^n \ \Rightarrow$ $(\nabla f(x) - \nabla f(y))^T(x-y) \geq 0 \ \forall \ x, y \in \mathbb{R}^n$

For all $x, y \in \mathbb{R}^n$, the subdifferential inequality gives

$$\begin{aligned} f(y) &\geq f(x) + \nabla f(x)^T(y-x) \\ f(x) &\geq f(y) + \nabla f(y)^T(x-y) \end{aligned}$$

Adding these two inequalities and canceling like terms from each side gives

$$0 \geq \nabla f(x)^T(y-x) + \nabla f(y)^T(x-y)$$

or equivalently

$$0 \leq (\nabla f(x) - \nabla f(y))^T(x-y) \ .$$

$(\nabla f(x) - \nabla f(y))^T(x-y) \geq 0 \;\forall\; x, y \in \mathbb{R}^n \;\Rightarrow$ $f(y) \geq f(x) + \nabla f(x)^T(y-x) \;\forall x, y \in \mathbb{R}^n$

Let $x, y \in \mathbb{R}^n$. By the MVT there exists $0 < \lambda < 1$ such that

$$f(y) - f(x) = \nabla f(x_\lambda)^T(y-x)$$

where $x_\lambda := \lambda y + (1-\lambda)x$. Then

$$\begin{aligned} 0 &\leq [\nabla f(x_\lambda) - \nabla f(x)]^T(x_\lambda - x) \\ &= \lambda[\nabla f(x_\lambda) - \nabla f(x)]^T(y-x) \\ &= \lambda[f(y) - f(x) - \nabla f(x)^T(y-x)]. \end{aligned}$$

Hence $f(y) \geq f(x) + \nabla f(x)^T(y-x)$.

$f(y) \geq f(x) + \nabla f(x)^T(y-x)\ \forall x, y \in \mathbb{R}^n \ \Rightarrow\ f$ is convex

Let $x, y \in \mathbb{R}^n$ and set

$$\alpha := \max_{\lambda \in [0,1]} \varphi(\lambda) := [f(\lambda y + (1-\lambda)x) - (\lambda f(y) + (1-\lambda) f(x))].$$

$f(y) \geq f(x) + \nabla f(x)^T(y-x)\ \forall x, y \in \mathbb{R}^n \;\Rightarrow\; f$ is convex

Let $x, y \in \mathbb{R}^n$ and set

$$\alpha := \max_{\lambda \in [0,1]} \varphi(\lambda) := [f(\lambda y + (1-\lambda)x) - (\lambda f(y) + (1-\lambda) f(x))].$$

Show $\alpha \leq 0$.

$f(y) \geq f(x) + \nabla f(x)^T(y-x)\ \forall x, y \in \mathbb{R}^n \ \Rightarrow\ f$ is convex

Let $x, y \in \mathbb{R}^n$ and set

$$\alpha := \max_{\lambda \in [0,1]} \varphi(\lambda) := [f(\lambda y + (1-\lambda)x) - (\lambda f(y) + (1-\lambda) f(x))].$$

Show $\alpha \leq 0$.
$[0,1]$ is compact and φ continuous, so $\exists\ \lambda \in [0,1]$ such that $\varphi(\lambda) = \alpha$.

$f(y) \geq f(x) + \nabla f(x)^T(y-x) \ \forall x, y \in \mathbb{R}^n \ \Rightarrow \ f$ is convex

Let $x, y \in \mathbb{R}^n$ and set

$$\alpha := \max_{\lambda \in [0,1]} \varphi(\lambda) := [f(\lambda y + (1-\lambda)x) - (\lambda f(y) + (1-\lambda)f(x))].$$

Show $\alpha \leq 0$.
$[0,1]$ is compact and φ continuous, so $\exists\ \lambda \in [0,1]$ such that $\varphi(\lambda) = \alpha$.
If λ equals 0 or 1, we are done, so assume $0 < \lambda < 1$.

$f(y) \geq f(x) + \nabla f(x)^T(y-x)\ \forall x, y \in \mathbb{R}^n \;\Rightarrow\; f$ is convex

Let $x, y \in \mathbb{R}^n$ and set

$$\alpha := \max_{\lambda \in [0,1]} \varphi(\lambda) := [f(\lambda y + (1-\lambda)x) - (\lambda f(y) + (1-\lambda)f(x))].$$

Show $\alpha \leq 0$.
$[0,1]$ is compact and φ continuous, so $\exists\ \lambda \in [0,1]$ such that $\varphi(\lambda) = \alpha$.
If λ equals 0 or 1, we are done, so assume $0 < \lambda < 1$. Then

$$0 = \varphi'(\lambda) = \nabla f(x_\lambda)^T(y-x) + f(x) - f(y)$$

where $x_\lambda = x + \lambda(y-x)$, or equivalently

$$\lambda(f(y) - f(x)) = -\nabla f(x_\lambda)^T(x - x_\lambda).$$

$f(y) \geq f(x) + \nabla f(x)^T(y-x)\ \forall x, y \in \mathbb{R}^n \ \Rightarrow\ f$ is convex

Let $x, y \in \mathbb{R}^n$ and set

$$\alpha := \max_{\lambda \in [0,1]} \varphi(\lambda) := [f(\lambda y + (1-\lambda)x) - (\lambda f(y) + (1-\lambda) f(x))].$$

Show $\alpha \leq 0$.
$[0,1]$ is compact and φ continuous, so $\exists\ \lambda \in [0,1]$ such that $\varphi(\lambda) = \alpha$.
If λ equals 0 or 1, we are done, so assume $0 < \lambda < 1$. Then

$$0 = \varphi'(\lambda) = \nabla f(x_\lambda)^T(y-x) + f(x) - f(y)$$

where $x_\lambda = x + \lambda(y-x)$, or equivalently

$$\lambda(f(y) - f(x)) = -\nabla f(x_\lambda)^T(x - x_\lambda).$$

Hence, by the subdifferential inequality,

$$\begin{aligned} \alpha &= f(x_\lambda) - (f(x) + \lambda(f(y) - f(x))) \\ &= f(x_\lambda) + \nabla f(x_\lambda)^T(x - x_\lambda) - f(x) \\ &\leq 0 \end{aligned}$$

f convex $\Rightarrow \nabla^2 f(x)$ PSD on $\mathbb{R}^n$

If f is convex,

$$f(x+td) \geq f(x) + t\nabla f(x)^T d \qquad \forall\ t \in \mathbb{R},\ x, d \in \mathbb{R}^n.$$

f convex $\Rightarrow \nabla^2 f(x)$ PSD on $\mathbb{R}^n$

If f is convex,

$$f(x+td) \geq f(x) + t\nabla f(x)^T d \qquad \forall\ t \in \mathbb{R},\ \ x, d \in \mathbb{R}^n.$$

So

$$f(x) + t\nabla f(x)^T d + \frac{t^2}{2} d^T \nabla^2 f(x) d + o(t^2) \geq f(x) + t\nabla f(x)^T d$$

f convex $\Rightarrow \nabla^2 f(x)$ PSD on $\mathbb{R}^n$

If f is convex,

$$f(x+td) \geq f(x) + t\nabla f(x)^T d \qquad \forall\ t \in \mathbb{R},\ x, d \in \mathbb{R}^n.$$

So

$$f(x) + t\nabla f(x)^T d + \frac{t^2}{2} d^T \nabla^2 f(x) d + o(t^2) \geq f(x) + t\nabla f(x)^T d$$

or equivalently

$$\frac{1}{2} d^T \nabla^2 f(x) d + \frac{o(t^2)}{t^2} \geq 0.$$

f convex $\Rightarrow \nabla^2 f(x)$ PSD on $\mathbb{R}^n$

If f is convex,

$$f(x+td) \geq f(x) + t\nabla f(x)^T d \qquad \forall\ t \in \mathbb{R},\ x, d \in \mathbb{R}^n.$$

So

$$f(x) + t\nabla f(x)^T d + \frac{t^2}{2} d^T \nabla^2 f(x) d + o(t^2) \geq f(x) + t\nabla f(x)^T d$$

or equivalently

$$\frac{1}{2} d^T \nabla^2 f(x) d + \frac{o(t^2)}{t^2} \geq 0.$$

Letting $t \to 0$ yields the inequality

$$d^T \nabla^2 f(x) d \geq 0.$$

f convex $\Leftarrow \nabla^2 f(x)$ PSD on $\mathbb{R}^n$

For $x, y \in \mathbb{R}^n$, the MVT implies there exists $\lambda \in (0,1)$ such that

$$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(x_\lambda)(y - x)$$

where $x_\lambda = \lambda y + (1 - \lambda)_x$.

f convex $\Leftarrow \nabla^2 f(x)$ PSD on $\mathbb{R}^n$

For $x, y \in \mathbb{R}^n$, the MVT implies there exists $\lambda \in (0,1)$ such that

$$f(y) = f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(x_\lambda)(y-x)$$

where $x_\lambda = \lambda y + (1-\lambda)_x$.
Hence

$$f(y) \geq f(x) + \nabla f(x)^T(y-x)$$

since $\nabla^2 f(x_\lambda)$ is positive semi-definite.

f convex $\Leftarrow \nabla^2 f(x)$ PSD on $\mathbb{R}^n$

For $x, y \in \mathbb{R}^n$, the MVT implies there exists $\lambda \in (0,1)$ such that

$$f(y) = f(x) + \nabla f(x)^T(y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(x_\lambda)(y-x)$$

where $x_\lambda = \lambda y + (1-\lambda)_x$.
Hence

$$f(y) \geq f(x) + \nabla f(x)^T(y-x)$$

since $\nabla^2 f(x_\lambda)$ is positive semi-definite.
Therefore, f is convex.

Checking for Positive Definiteness

Theorem: *Let $H \in \mathbb{R}^{n \times n}$ be symmetric. We define the kth principal minor of H, denoted $\Delta_k(H)$, to be the determinant of the upper left-hand $k \times k$ submatrix of H. Then H is positive definite if and only if $\Delta_k(H) > 0$ for $k = 1, 2, \ldots, n$.*

Checking for Positive Definiteness

Example:

$$H = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 5 & 1 \\ -1 & 1 & 4 \end{bmatrix}.$$

We have

$$\Delta_1(H) = 1, \quad \Delta_2(H) = \begin{vmatrix} 1 & 1 \\ 1 & 5 \end{vmatrix} = 4, \quad \Delta_3(H) = \det(H) = 8.$$

Therefore, H is positive definite.

Sublinear Functionals

Definition: *Let* $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. *We say that h is positively homogeneous if*

$$h(\lambda x) = \lambda h(x) \quad \text{for all } x \in \mathbb{R} \text{ and } \lambda > 0.$$

We say that h is subadditive if

$$h(x + y) \leq h(x) + h(y) \quad \text{for all } x, y \in \mathbb{R}.$$

Finally, we say that h is sublinear if it is both subadditive and positively homogeneous.

Sublinear Functionals

Definition: *Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. We say that h is positively homogeneous if*

$$h(\lambda x) = \lambda h(x) \quad \textit{for all } x \in \mathbb{R} \textit{ and } \lambda > 0.$$

We say that h is subadditive if

$$h(x + y) \leq h(x) + h(y) \quad \textit{for all } x, y \in \mathbb{R}.$$

Finally, we say that h is sublinear if it is both subadditive and positively homogeneous.

Note: Sublinear functionals can be thought of as a generalization of the notion of a norm. In particular, every norm is a sublinear functional.

Sublinear Functionals

Sublinear functionals form a very rich and important class of functionals with a surprisingly beautiful structure.

Sublinear Functionals

Sublinear functionals form a very rich and important class of functionals with a surprisingly beautiful structure.

Lemma: *Every sublinear functional is a convex function.*

Sublinear Functionals

Sublinear functionals form a very rich and important class of functionals with a surprisingly beautiful structure.

Lemma: *Every sublinear functional is a convex function.*

Proof:

$$\begin{aligned} h(\lambda x + (1-\lambda)y) &\leq h(\lambda x) + h(1-\lambda)y) \\ &= \lambda h(x) + (1-\lambda)h(y). \end{aligned}$$

Sublinear Functionals

Sublinear functionals form a very rich and important class of functionals with a surprisingly beautiful structure.

Lemma: *Every sublinear functional is a convex function.*

Proof:

$$\begin{aligned} h(\lambda x+(1-\lambda)y) &\leq h(\lambda x)+h(1-\lambda)y) \\ &= \lambda h(x)+(1-\lambda)h(y). \end{aligned}$$

Lemma: *A function is sublinear if and only if its epigraph is a convex cone.*

Directional Derivatives of Convex Functions are Sublinear

Theorem: *Let* $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ *be a convex function.*

Directional Derivatives of Convex Functions are Sublinear

Theorem: *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a convex function.*

Then at every point $x \in \mathrm{dom}\,(f)$ *the directional derivative* $f'(x; d)$ *is a sublinear function of the* d *argument.*

Directional Derivatives of Convex Functions are Sublinear

Theorem: *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a convex function.*

Then at every point $x \in \operatorname{dom}(f)$ *the directional derivative* $f'(x; d)$ *is a sublinear function of the* d *argument.*

That is, the function $f'(x; \cdot) : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is sublinear.*

Directional Derivatives of Convex Functions are Sublinear

Theorem: *Let* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *be a convex function.*

Then at every point $x \in \mathrm{dom}\,(f)$ *the directional derivative* $f'(x; d)$ *is a sublinear function of the* d *argument.*

That is, the function $f'(x; \cdot) : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is sublinear.*

Thus, in particular, the function $f'(x; \cdot)$ *is a convex function.*

Directional Derivatives of Convex Functions are Sublinear

Proof: (Positive Homogeneity) Let $x \in \operatorname{dom}(f)$, $d \in \mathbb{R}^n$, and $\lambda > 0$. Then

$$f'(x; \lambda d) = \lim_{t \downarrow 0} \frac{f(x + t\lambda d) - f(x)}{t}$$

Directional Derivatives of Convex Functions are Sublinear

Proof: (Positive Homogeneity) Let $x \in \mathrm{dom}\,(f)$, $d \in \mathbb{R}^n$, and $\lambda > 0$. Then

$$
\begin{aligned}
f'(x;\lambda d) &= \lim_{t \downarrow 0} \frac{f(x + t\lambda d) - f(x)}{t} \\
&= \lim_{t \downarrow 0} \lambda \frac{f(x + t\lambda d) - f(x)}{\lambda t}
\end{aligned}
$$

Directional Derivatives of Convex Functions are Sublinear

Proof: (Positive Homogeneity) Let $x \in \operatorname{dom}(f)$, $d \in \mathbb{R}^n$, and $\lambda > 0$. Then

$$\begin{aligned} f'(x;\lambda d) &= \lim_{t\downarrow 0} \frac{f(x+t\lambda d)-f(x)}{t} \\ &= \lim_{t\downarrow 0} \lambda \frac{f(x+t\lambda d)-f(x)}{\lambda t} \\ &= \lambda \lim_{(\lambda t)\downarrow 0} \frac{f(x+(t\lambda)d)-f(x)}{(\lambda t)} \end{aligned}$$

Directional Derivatives of Convex Functions are Sublinear

Proof: (Positive Homogeneity) Let $x \in \mathrm{dom}\,(f)$, $d \in \mathbb{R}^n$, and $\lambda > 0$. Then

$$\begin{aligned} f'(x;\lambda d) &= \lim_{t\downarrow 0} \frac{f(x+t\lambda d)-f(x)}{t} \\ &= \lim_{t\downarrow 0} \lambda \frac{f(x+t\lambda d)-f(x)}{\lambda t} \\ &= \lambda \lim_{(\lambda t)\downarrow 0} \frac{f(x+(t\lambda)d)-f(x)}{(\lambda t)} \\ &= \lambda f'(x;d). \end{aligned}$$

Directional Derivatives of Convex Functions are Sublinear

(Subadditivity) let $d_1, d_2 \in \mathbb{R}^n$, Then

$$f'(x; d_1 + d_2) = \lim_{t \downarrow 0} \frac{f(x + t(d_1 + d_2)) - f(x)}{t}$$

Directional Derivatives of Convex Functions are Sublinear

(Subadditivity) let $d_1, d_2 \in \mathbb{R}^n$, Then

$$\begin{aligned} f'(x; d_1 + d_2) &= \lim_{t \downarrow 0} \frac{f(x + t(d_1 + d_2)) - f(x)}{t} \\ &= \lim_{t \downarrow 0} \frac{f(\frac{1}{2}(x + 2td_1) + \frac{1}{2}(x + 2td_2)) - f(x)}{t} \end{aligned}$$

Directional Derivatives of Convex Functions are Sublinear

(Subadditivity) let $d_1, d_2 \in \mathbb{R}^n$, Then

$$\begin{aligned} f'(x; d_1 + d_2) &= \lim_{t \downarrow 0} \frac{f(x + t(d_1 + d_2)) - f(x)}{t} \\ &= \lim_{t \downarrow 0} \frac{f(\frac{1}{2}(x + 2td_1) + \frac{1}{2}(x + 2td_2)) - f(x)}{t} \\ &\leq \lim_{t \downarrow 0} \frac{\frac{1}{2}f(x + 2td_1) + \frac{1}{2}f(x + 2td_2) - f(x)}{t} \end{aligned}$$

Directional Derivatives of Convex Functions are Sublinear

(Subadditivity) let $d_1, d_2 \in \mathbb{R}^n$, Then

$$
\begin{aligned}
f'(x; d_1 + d_2) &= \lim_{t \downarrow 0} \frac{f(x + t(d_1 + d_2)) - f(x)}{t} \\
&= \lim_{t \downarrow 0} \frac{f(\frac{1}{2}(x + 2td_1) + \frac{1}{2}(x + 2td_2)) - f(x)}{t} \\
&\leq \lim_{t \downarrow 0} \frac{\frac{1}{2} f(x + 2td_1) + \frac{1}{2} f(x + 2td_2) - f(x)}{t} \\
&\leq \lim_{t \downarrow 0} \frac{\frac{1}{2}(f(x + 2td_1) - f(x)) + \frac{1}{2}(f(x + 2td_2) - f(x))}{t}
\end{aligned}
$$

Directional Derivatives of Convex Functions are Sublinear

(Subadditivity) let $d_1, d_2 \in \mathbb{R}^n$, Then

$$
\begin{aligned}
f'(x; d_1 + d_2) &= \lim_{t \downarrow 0} \frac{f(x + t(d_1 + d_2)) - f(x)}{t} \\
&= \lim_{t \downarrow 0} \frac{f(\frac{1}{2}(x + 2td_1) + \frac{1}{2}(x + 2td_2)) - f(x)}{t} \\
&\leq \lim_{t \downarrow 0} \frac{\frac{1}{2} f(x + 2td_1) + \frac{1}{2} f(x + 2td_2) - f(x)}{t} \\
&\leq \lim_{t \downarrow 0} \frac{\frac{1}{2}(f(x + 2td_1) - f(x)) + \frac{1}{2}(f(x + 2td_2) - f(x))}{t} \\
&= \lim_{t \downarrow 0} \frac{f(x + 2td_1) - f(x)}{2t} + \lim_{t \downarrow 0} \frac{f(x + 2td_2) - f(x)}{2t}
\end{aligned}
$$

Directional Derivatives of Convex Functions are Sublinear

(Subadditivity) let $d_1, d_2 \in \mathbb{R}^n$, Then

$$\begin{aligned} f'(x; d_1 + d_2) &= \lim_{t \downarrow 0} \frac{f(x + t(d_1 + d_2)) - f(x)}{t} \\ &= \lim_{t \downarrow 0} \frac{f(\frac{1}{2}(x + 2td_1) + \frac{1}{2}(x + 2td_2)) - f(x)}{t} \\ &\leq \lim_{t \downarrow 0} \frac{\frac{1}{2} f(x + 2td_1) + \frac{1}{2} f(x + 2td_2) - f(x)}{t} \\ &\leq \lim_{t \downarrow 0} \frac{\frac{1}{2}(f(x + 2td_1) - f(x)) + \frac{1}{2}(f(x + 2td_2) - f(x))}{t} \\ &= \lim_{t \downarrow 0} \frac{f(x + 2td_1) - f(x)}{2t} + \lim_{t \downarrow 0} \frac{f(x + 2td_2) - f(x)}{2t} \\ &= f'(x; d_1) + f'(x; d_2). \end{aligned}$$

Examples of Convex Functions

Set $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \; | 0 < x_i, \;\; i = 1, \ldots, n\}$ and $\mathbb{S}^n_{++}$ equal to the set on $n \times n$ symmetric positive definite matrices.

Examples of Convex Functions

Set $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \mid 0 < x_i,\ i = 1, \ldots, n\}$ and $\mathbb{S}^n_{++}$ equal to the set on $n \times n$ symmetric positive definite matrices.

$$f(x) = \|x\|$$

Examples of Convex Functions

Set $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \;|0 < x_i,\; i = 1, \ldots, n\}$ and $\mathbb{S}^n_{++}$ equal to the set on $n \times n$ symmetric positive definite matrices.

$$\begin{aligned} f(x) &= \|x\| \\ f(x) &= \max\{x_1, x_2, \ldots, x_n\} \end{aligned}$$

Examples of Convex Functions

Set $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \mid 0 < x_i,\ i = 1, \ldots, n\}$ and $\mathbb{S}^n_{++}$ equal to the set on $n \times n$ symmetric positive definite matrices.

$$\begin{aligned} f(x) &= \|x\| \\ f(x) &= \max\{x_1, x_2, \ldots, x_n\} \\ f(x) &= \frac{1}{2}x^T Hx + g^T x + \alpha \qquad H \in \mathbb{S}^n_{++} \end{aligned}$$

Examples of Convex Functions

Set $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \mid 0 < x_i,\ i = 1, \ldots, n\}$ and $\mathbb{S}^n_{++}$ equal to the set on $n \times n$ symmetric positive definite matrices.

$$
\begin{aligned}
f(x) &= \|x\| \\
f(x) &= \max\{x_1, x_2, \ldots, x_n\} \\
f(x) &= \frac{1}{2}x^T Hx + g^T x + \alpha \qquad H \in \mathbb{S}^n_{++} \\
f(x) &= \log(e^{x_1} + e^{x_2} + \cdots + e^{x_n})
\end{aligned}
$$

Examples of Convex Functions

Set $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \;|0 < x_i,\; i = 1, \ldots, n\}$ and $\mathbb{S}^n_{++}$ equal to the set on $n \times n$ symmetric positive definite matrices.

$$
\begin{aligned}
f(x) &= \|x\| \\
f(x) &= \max\{x_1, x_2, \ldots, x_n\} \\
f(x) &= \frac{1}{2}x^T Hx + g^T x + \alpha \qquad H \in \mathbb{S}^n_{++} \\
f(x) &= \log(e^{x_1} + e^{x_2} + \cdots + e^{x_n}) \\
f(x) &= -\left(\prod_{j=1}^{n} x_i\right)^{1/n} \qquad x \in \mathbb{R}^n_{++}
\end{aligned}
$$

Examples of Convex Functions

Set $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \mid 0 < x_i,\ i = 1, \ldots, n\}$ and $\mathbb{S}^n_{++}$ equal to the set on $n \times n$ symmetric positive definite matrices.

$$
\begin{aligned}
f(x) &= \|x\| \\
f(x) &= \max\{x_1, x_2, \ldots, x_n\} \\
f(x) &= \frac{1}{2}x^T Hx + g^T x + \alpha \qquad H \in \mathbb{S}^n_{++} \\
f(x) &= \log(e^{x_1} + e^{x_2} + \cdots + e^{x_n}) \\
f(x) &= -\left(\prod_{j=1}^n x_i\right)^{1/n} \qquad x \in \mathbb{R}^n_{++} \\
f(x) &= -\log(x_1 x_2 \ldots x_n) \qquad x \in \mathbb{R}^n_{++}
\end{aligned}
$$

Examples of Convex Functions

Set $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \,|0 < x_i,\ \ i = 1, \ldots, n\}$ and $\mathbb{S}^n_{++}$ equal to the set on $n \times n$ symmetric positive definite matrices.

$$
\begin{aligned}
f(x) &= \|x\| \\
f(x) &= \max\{x_1, x_2, \ldots, x_n\} \\
f(x) &= \frac{1}{2}x^T Hx + g^T x + \alpha \qquad H \in \mathbb{S}^n_{++} \\
f(x) &= \log(e^{x_1} + e^{x_2} + \cdots + e^{x_n}) \\
f(x) &= -\left(\prod_{j=1}^{n} x_i\right)^{1/n} \qquad x \in \mathbb{R}^n_{++} \\
f(x) &= -\log(x_1 x_2 \ldots x_n) \qquad x \in \mathbb{R}^n_{++} \\
f(x) &= -\log\det(H)) \qquad H \in \mathbb{S}^n_{++}
\end{aligned}
$$