Advanced Mathematical Programming IE417

Lecture 4

Dr. Ted Ralphs

Reading for This Lecture

- Chapter 3, Sections 1-3

Convex Functions

Convex Functions

Definition 1. *Let S be a nonempty convex set on $\mathbb{R}^n$. Then the function $f : S \to \mathbb{R}$ is said to be* convex *on S if*

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

for each $x_1, x_2 \in S$ and $\lambda \in (0,1)$.

- *Strictly convex* means the inequality is strict.
- *(Strictly) concave* is defined analogously.
- Can a function be concave and convex?

Properties of Convex Functions

- A positive combination of convex functions is convex.
- Suppose $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is concave and let $S = \{x : g(x) > 0\}$. Then if $f : S \rightarrow \mathbb{R}$ is defined by $f(x) = 1/g(x), f$ is convex.
- $f : S \rightarrow \mathbb{R}$ is concave if and only if $-f$ is convex.
- Let S be a nonempty convex set on $\mathbb{R}^n$ and let $f : S \rightarrow \mathbb{R}$ be convex. Then the *level set* $S_\alpha = \{x \in S : f(x) \leq \alpha\}$, where $\alpha \in \mathbb{R}$, is a convex set.

Continuity of Convex Functions

Theorem 1. *Let S be a nonempty convex set on $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$ be convex. Then f is continuous on the interior of S.*

Proof Idea:

Directional Derivative

Definition 2. *Let S be a nonempty set on $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$ be convex. For $x^* \in S$ and $d \in \mathbb{R}^n$ such that $x^* + \lambda d \in S$ for $\lambda > 0$ sufficiently small. The* directional derivative *at x^* along d is given by the following limit, if it exists:*

$$f(x^*; d) = \lim_{\lambda \to 0} (f(x^* + \lambda d) - f(x^*))/\lambda$$

- Directional derivatives always exist for convex functions that are defined on all of $\mathbb{R}^n$.

The Epigraph

Definition 3. *Let S be a nonempty set on $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$. The* epigraph *of f is a subset of $\mathbb{R}^{n+1}$ defined by*

$$\text{epi} f = \{(x, y) : x \in S, y \in \mathbb{R}, y \geq f(x)\}$$

Theorem 2. *Let S be a nonempty convex set on $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$. f is convex if and only if* $\text{epi} f$ *is a convex set.*

Subgradients

Definition 4. *Let S be a nonempty set on $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$. Then ξ is called a* subgradient *of $x^* \in S$ if*

$$f(x) \geq f(x^*) + \xi^T(x - x^*)$$

- The collection of all subgradients is a convex set and is called the *subdifferential*.

More Subgradient Results

Theorem 3. *Let* S *be a nonempty convex set on* $\mathbb{R}^n$ *and let* $f : S \longrightarrow \mathbb{R}$ *be convex. Then for each* $x^* \in \text{int } S$*, there exists a vector* ξ *such that the hyperplane*

$$H = \{(x, y) : y = f(x^*) + \xi^T(x - x^*)\}$$

supports $\text{epi} f$ *at* $[x^*, f(x^*)]$*.*

Theorem 4. *Let* S *be a nonempty convex set on* $\mathbb{R}^n$ *and let* $f : S \longrightarrow \mathbb{R}$*. If for each point* $x^* \in \text{int } S$ *there exists a subgradient vector, then* f *is convex on* $\text{int } S$*.*

The Gradient

Definition 5. *Let S be a nonempty set on $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$. Then, f is said to be* differentiable *at $x^* \in \text{int } S$ if there exists a vector $\nabla f(x^*)$ called the* gradient*, and a function $\alpha : \mathbb{R}^n \to \mathbb{R}$ such that*

$$f(x) = f(x^*) + \nabla f(x^*)^T (x - x^*) + \|x - x^*\| \alpha(x^*; x - x^*)$$

for each $x \in S$ where $\lim_{x \to x^*} \alpha(x^*; x - x^*) = 0$

The Subgradient and the Gradient

Lemma 1. *Let S be a nonempty convex set in $\mathbb{R}^n$ and let $f : S \rightarrow \mathbb{R}$ be convex. Suppose that f is differentiable at $x^* \in \text{int } S$. Then the only subgradient at x^* is the gradient.*

Theorem 5. *Let S be a nonempty convex set in $\mathbb{R}^n$ and let $f : S \rightarrow \mathbb{R}$ be differentiable on S. Then f is convex if and only if for any $x^* \in S$ we have*

$$f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*) \; \forall x \in S$$

Twice Differentiable Functions

Definition 6. *Let S be a nonempty set in $\mathbb{R}^n$ and let $f : S \rightarrow \mathbb{R}$. Then f is said to be* twice differentiable *at $x^* \in \text{int } S$ if there exists a vector $\nabla f(x^*)$ and an $n \times n$ symmetric matrix $H(x^*)$, called the* Hessian matrix *and a function $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ such that*

$$f(x) = f(x^*) + \nabla f(x^*)^T(x - x^*) + \frac{1}{2}(x - x^*)^T H(x^*)(x - x^*)$$
$$+\|x - x^*\|\alpha(x^*; x - x^*)$$

for each $x \in S$ where $\lim_{x \rightarrow x^*} \alpha(x^*; x - x^*) = 0$.

Positive Semidefinite Matrices

Definition 7. *A matrix* $M \in \mathbb{R}^{n \times n}$ *is said to be* positive semidefinite *if* $x^T M x \geq 0 \; \forall x \in \mathbb{R}^n$.

Theorem 6. *Let* S *be a nonempty open convex set in* $\mathbb{R}^n$ *and let* $f : S \to \mathbb{R}$ *be twice differentiable on* S*. Then* f *is convex if and only if the Hessian matrix is positive semidefinite at each point in* S*.*

Strict Convexity

Theorem 7. *Let S be a nonempty convex set on $\mathbb{R}^n$ and let $f : S \to \mathbb{R}$ be twice differentiable on S. If the Hessian matrix is positive definite at each point in S, then f is strictly convex.*

- The converse is not necessarily true.
- Example: $f(x) = x^4$.

Infinitely Differentiable Functions

Theorem 8. *Let* S *be a nonempty open convex set in* $\mathbb{R}$ *and let* $f : S \to \mathbb{R}$ *be infinitely differentiable on* S. *Then* f *is strictly convex on* f *if and only if, for each* $x^* \in S$, *there exists an even* n *such that* $f^{(n)}(x^*) > 0$ *while* $f^{(j)}(x^*) = 0 \;\; \forall j < n$.

- *This fixes the problem with the previous example.*

Recognizing PSD Matrices

Lemma 2. *If $H \in \mathbb{R}^{2\times 2}$ has elements h_{ij}, then H is positive semidefinite iff $h_{ii} \geq 0$ and $h_{11}h_{22} - h_{12}^2 \geq 0$. H is positive definite if these inequalities are strict.*

Theorem 9. *Let $H \in \mathbb{R}^{n\times n}$ be a symmetric matrix with elements h_{ij}.*

- *If $h_{ii} \leq 0$ for some i, then H is not positive definite; and if $h_{ii} < 0$ for some i, then H is not positive semidefinite.*
- *If $h_{ii} = 0$ for some i, then we must have $h_{ij} = h_{ji} = 0$ for all j or else H is not positive semidefinite.*

Algorithm for Recognizing PSD Matrices

Theorem 10. *If $n = 1$, H is positive semidefinite (positive definite) if and only if $h_{11} \geq 0$ ($h_{11} > 0$). Otherwise, if $n \geq 2$, let*

$$H = \begin{bmatrix} h_{11} & q^T \\ q & G \end{bmatrix}$$

where $q = 0$ if $h_{11} = 0$, and otherwise $h_{11} > 0$. Using elementary row operations, reduce H to the following form:

$$\begin{bmatrix} h_{11} & q^T \\ 0 & G_{new} \end{bmatrix}.$$

Then H is positive semidefinite if and only if G_{new} is positive semidefinite. Furthermore, if $h_{11} > 0$, then H is positive definite if and only if G_{new} is positive definite.