# Convex Analysis - Convex Functions

## Nonlinear Programming

Fatih Cavdur

to accompany
Nonlinear Programming
by Mokhtar S. Bazara, Hanif D. Sherali and C. M. Shetty

## Definitions and Basic Properties

### Definition

Let $f : S \to R$, where $S$ is a nonempty convex set in $R^n$. The function is said to be convex on $S$ if

$$f(\lambda \mathbf{x_1} + (1-\lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x_1}) + (1-\lambda) f(\mathbf{x}_2)$$

for each $\mathbf{x}_1, \mathbf{x_2} \in S$ and for each $\lambda \in (0,1)$.
The function $f$ is called strictly convex on $S$ if the above inequality is true as a strict inequality for each distinct $\mathbf{x}_1$ and $\mathbf{x}_2$ in $S$ and for each $\lambda \in (0,1)$.
The function $f : S \to R$ is called concave (strictly concave) on $S$ if $-f$ is convex (strictly convex) on $S$.
A function is both convex and concave if and only if it is affine.

## Definitions and Basic Properties

### Lemma

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$ be a convex function. Then the level set $S_\alpha = \{\mathbf{x} \in S : f(\mathbf{x} \leq \alpha\}$, where $\alpha$ is a real number, is a convex set.

## Definitions and Basic Properties

### Proof

Let $\mathbf{x}_1, \mathbf{x}_2 \in S_\alpha$. Thus, $\mathbf{x}_1, \mathbf{x}_2 \in S$ and $f(\mathbf{x}_1) \leq \alpha$ and $f(\mathbf{x}_2) \leq \alpha$. Now let $\lambda \in (0, 1)$ and $\mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$. By the convexity of $S$, we have that $\mathbf{x} \in S$. Furthermore, by the convexity of $f$,

$$f(\mathbf{x}) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2) \leq \lambda\alpha + (1 - \lambda)\alpha = \alpha$$

Hence, $\mathbf{x} \in S_\alpha$, and therefore, $S_\alpha$ is convex.

## Continuity of Convex Functions

An important property of convex and concave functions is that they are continuous on the interior on their domain.

### Theorem

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$ be convex. Then $f$ is continuous on the interior of $S$.

## Directional Derivative of Convex Functions

### Definition

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$. Let $\bar{\mathbf{x}} \in S$ and $\mathbf{d}$ be a nonzero vector such that such that $\bar{\mathbf{x}} + \lambda\mathbf{d} \in S$ for $\lambda > 0$ and sufficiently small. The directional derivative of $f$ at $\bar{\mathbf{x}}$ along the vector $\mathbf{d}$, denoted by $f'(\bar{\mathbf{x}}; \mathbf{d})$, is given by the following limit if it exists:

$$f'(\bar{\mathbf{x}}; \mathbf{d}) = \lim_{\lambda \to 0^+} \frac{f(\bar{\mathbf{x}}; \mathbf{d}) - f(\bar{\mathbf{x}})}{\lambda}$$

## Directional Derivative of Convex Functions

### Lemma

Let $f : R^n \to R$ be a convex function. Consider any point $\bar{\mathbf{x}} \in R^n$ and a nonzero direction $\mathbf{d} \in R^n$. Then the directional derivative $f'(\bar{\mathbf{x}}; \mathbf{d})$, of $f$ at $\bar{\mathbf{x}}$ in the direction $\mathbf{d}$, exists.

## Directional Derivative of Convex Functions

### Lemma

Let $f : R^n \to R$ be a convex function. Consider any point $\bar{\mathbf{x}} \in R^n$ and a nonzero direction $\mathbf{d} \in R^n$. Then the directional derivative $f'(\bar{\mathbf{x}}; \mathbf{d})$, of $f$ at $\bar{\mathbf{x}}$ in the direction $\mathbf{d}$, exists.

## Epigraph and Hypograph of a Function

### Definition

A function $f$ on $S$ can be fully described by the set $\{[\mathbf{x}, f(\mathbf{x})] : \mathbf{x} \in S\} \subset R^{n+1}$, which is referred to as the graph of the function. One can construct two sets that are related to the graph of $f$:

1. the epigraph, which consists of points above the graph of $f$
2. the hypograph, which consists of points below the graph of $f$

## Epigraph and Hypograph of a Function

### Definition

Let $S$ be a nonempty set in $R^n$, and let $f : S \to R$. The epigraph of $f$, denoted by epi $f$, is a subset of $R^{n+1}$ defined by

$$\{(\mathbf{x}, y) : \mathbf{x} \in S, y \in R, y \geq f(\mathbf{x})\}$$

The hypograph of $f$, denoted by hyp $f$, is a subset of $R^{n+1}$ defined by

$$\{(\mathbf{x}, y) : \mathbf{x} \in S, y \in R, y \leq f(\mathbf{x})\}$$

## Epigraph and Hypograph of a Function

### Theorem

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$. Then $f$ is convex if and only if epi $f$ is a convex set.

## Subgradient of a Function

### Definition

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$ be convex. Then $\xi$ is called a subgradient of $f$ at $\bar{\mathbf{x}}$ if

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$

Similarly, let $f : S \to R$ be concave. Then $\xi$ is called a subgradient of $f$ at $\bar{\mathbf{x}}$ if

$$f(\mathbf{x}) \leq f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$


## Subgradient of a Function

### Definition

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$ be convex. Then $\xi$ is called a subgradient of $f$ at $\bar{\mathbf{x}}$ if

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$

Similarly, let $f : S \to R$ be concave. Then $\xi$ is called a subgradient of $f$ at $\bar{\mathbf{x}}$ if

$$f(\mathbf{x}) \leq f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$

## Subgradient of a Function

### Definition

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$ be convex. Then for $\bar{\mathbf{x}} \in \text{int}(S)$, there exists a vector $\xi$ such that the hyperplane

$$H = \{(\mathbf{x}, y) : y = f(\mathbf{x}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}\}$$

supports epi $f$ at $[\bar{\mathbf{x}}, f(\bar{\mathbf{x}})]$. In particular,

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$

that is, $\xi$ is a subgradient of $f$ at $\bar{\mathbf{x}}$.

## Subgradient of a Function

### Corollary

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$ be strictly convex. Then for $\bar{\mathbf{x}} \in \text{int}(S)$, there exists a vector $\xi$ such that

$$f(\mathbf{x}) > f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$

## Subgradient of a Function

### Theorem

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$. Suppose that for each point $\bar{\mathbf{x}} \in \text{int}S$ there exists a subgradient vector $\xi$ such that

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$

Then, $f$ is convex on int$S$.

## Differentiable Functions

### Definition

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

$$f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}}^t)^t(\mathbf{x} - \bar{\mathbf{x}}) + ||\mathbf{x} - \bar{\mathbf{x}}||\alpha(\bar{\mathbf{x}}; \mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$

where $\lim_{\mathbf{x} \to \bar{\mathbf{x}}} \alpha(\bar{\mathbf{x}}; \mathbf{x} - \bar{\mathbf{x}}) = 0$. The function $f$ is said to be differentiable on the open set $S' \subseteq S$ if it is differentiable at each point in $S'$. This representation of $f$ is called a first-order Taylor series expansion of $f$ at (or about) the point $\bar{\mathbf{x}}$.

## Differentiable Functions

### Definition

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$ be convex. Suppose that $f$ is differentiable at $\bar{\mathbf{x}} \in \text{int}S$. Then the collection of subgradients of $f$ at $\bar{\mathbf{x}}$ is the singleton set $\{\nabla f(\bar{\mathbf{x}})\}$.

## Differentiable Functions

### Theorem

Let $S$ be a nonempty open convex set in $R^n$, and let $f : S \to R$ be differentiable on $S$. Then $f$ is convex if and only if for any $\bar{\mathbf{x}} \in S$, we have

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \in S$$

Similarly, $f$ is strictly convex if and only if, we have

$$f(\mathbf{x}) > f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^t(\mathbf{x} - \bar{\mathbf{x}}), \quad \forall \mathbf{x} \neq \bar{\mathbf{x}} \in S$$

# Differentiable Functions

### Theorem

Let $S$ be a nonempty open convex set in $R^n$, and let $f : S \to R$ be differentiable on $S$. Then $f$ is convex if and only if for each $\mathbf{x}_1, \mathbf{x_2} \in S$, we have

$$[\nabla f(\mathbf{x}_2 - \nabla \mathbf{x}_1)]^t(\mathbf{x}_2 - \mathbf{x}_1) \geq 0$$

Similarly, $f$ is strictly convex if and only if, for each distinct $\mathbf{x}_1, \mathbf{x}_2 \in S$, we have

$$[\nabla f(\mathbf{x}_2 - \nabla \mathbf{x}_1)]^t(\mathbf{x}_2 - \mathbf{x}_1) > 0$$

# Twice Differentiable Convex and Concave Functions

### Definition

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

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

for each $\mathbf{x} \in S$, where $\lim_{\mathbf{x}\to\bar{\mathbf{x}}} \alpha(\bar{\mathbf{x}}; \mathbf{x}-\bar{\mathbf{x}}) = 0$. The function $f$ is said to be twice differentiable on the open set $S' \subseteq S$ if it is twice differentiable at each point $S'$.

## Twice Differentiable Functions

We can write the foregoing representation as follows, which, without the remainder term, is known as a second-order (Taylor series) approximation at (or about) the point $\bar{\mathbf{x}}$.

$$
\begin{aligned}
f(\mathbf{x}) = f(\bar{\mathbf{x}}) &+ \sum_{j=1}^{n} f_j(\bar{\mathbf{x}})(x_j - \bar{x}_j) \\
&+ \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - \bar{x}_i)(x_j - \bar{x}_j) f_{ij}(\bar{\mathbf{x}}) \\
&+ \alpha(\bar{\mathbf{x}}; \mathbf{x} - \bar{\mathbf{x}})
\end{aligned}
$$

## Example

Let $f(x_1, x_2) = 2x_1 + 6x_2 - 2x_1^2 - 3x_2^2 + 4x_1x_2$. Then we have,

$$
\nabla f(\bar{\mathbf{x}}) = \begin{bmatrix} 2 - 4\bar{x}_1 + 4\bar{x}_2 \\ 6 - 6\bar{x}_2 + 4\bar{x}_1 \end{bmatrix} \quad \text{and} \quad \mathbf{H}(\bar{\mathbf{x}}) = \begin{bmatrix} -4 & 4 \\ 4 & -6 \end{bmatrix}
$$

For example, taking $\bar{\mathbf{x}} = (0, 0)^t$, the second-order expansion of this function is given by

$$
f(x_1, x_2) = (2, 6)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{2}(x_1, x_2)\begin{bmatrix} -4 & 4 \\ 4 & -6 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}
$$

Note that there is no remainder term here since the given function is quadratic, so the above expression is exact.

## Differentiable Functions

Let $S$ be a nonempty open convex set in $R^n$, and let $f : S \to R$ be twice differentiable on $S$. Then $f$ is convex if and only if the Hessian matrix is positive semidefinite at each point in $S$; that is, for any $\bar{x} \in S$, we have $\mathbf{x}^t\mathbf{H}(\bar{\mathbf{x}})\mathbf{x} \geq 0$ for all $\mathbf{x} \in R^n$.
A function is concave if and only if its Hessian matrix is negative semidefinite (NSD) everywhere in $S$; that is, for any $\bar{x} \in S$, we have $\mathbf{x}^t\mathbf{H}(\bar{\mathbf{x}})\mathbf{x} \leq 0$ for all $\mathbf{x} \in R^n$.
A matrix is that is neither PSD or NSD is called indefinite (ID).

## Differentiable Functions

### Theorem

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

# Differentiable Functions

### Theorem

Let $S$ be a nonempty open convex set in $R^n$, and let $f : S \to R$ be twice differentiable on $S$. If the Hessian matrix is positive definite (PD) at each point in $S$, $f$ is strictly convex. Conversely, if $f$ is strictly convex, the Hessian matrix is PSD at each point in $S$. However, if $f$ is strictly convex and quadratic, its Hessian is PD.

# Differentiable Functions

### Theorem

Let $S$ be a nonempty open convex set in $R^n$, and let $f : S \to R$ be infinitely differentiable on $S$. Then $f$ is strictly convex on $S$ if and only if for each $\bar{x} \in S$, there exists an even $n$ such that $f^{(n)}(\bar{x}) > 0$, while $f^{(j)}(\bar{x}) = 0$ for any $1 < j < n$.

## Differentiable Functions

### Theorem

Consider a function $f : R^n \rightarrow R$, and for any point $\bar{\mathbf{x}} \in R^n$ and a nonzero direction $\mathbf{d} \in R^n$, define $F_{(\bar{\mathbf{x}};\mathbf{d})} = f(\bar{\mathbf{x}} + \lambda \mathbf{d})$ as a function of $\lambda \in R$. Then $f$ is (strictly) convex if and only if $F_{(\bar{\mathbf{x}};\mathbf{d})}$ is (strictly) convex for all $\bar{\mathbf{x}}$ and $\mathbf{d} \neq 0$ in $R^n$.

## Minimizing a Convex Function

### Definition

Let $f : R^n \rightarrow R$ and consider the problem to minimize $f(\mathbf{x})$ subject to $\mathbf{x} \in S$.
A point $\mathbf{x} \in S$ is called a feasible solution to the problem.
If $\bar{\mathbf{x}} \in S$ and $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$ for each $\mathbf{x} \in S$, $\bar{\mathbf{x}}$ is called an optimal solution, a global optimal solution, or a solution to the problem.
The collections of optimal solutions are called alternative optimal solutions.
If $\bar{\mathbf{x}} \in S$ and there exists an $\epsilon$-neighborhood $N_\epsilon(\bar{\mathbf{x}})$ around $\bar{\mathbf{x}}$ such that $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$ for each $\mathbf{x} \in S \cap N_\epsilon(\bar{\mathbf{x}})$ , $\bar{\mathbf{x}}$ is called a local optimal solution.

## Minimizing a Convex Function

### Theorem

Let $S$ be a nonempty convex set in $R^n$, and let $f : S \to R$ be convex on $S$. Consider the problem to minimize $f(\mathbf{x})$ subject to $\mathbf{x} \in S$. Suppose that $\bar{\mathbf{x}} \in S$ is a local optimal solution to the problem. Then, $\bar{\mathbf{x}}$ is a global solution.
If either $\bar{\mathbf{x}}$ is a strict local minimum or $f$ is strictly convex, $\bar{\mathbf{x}}$ is the unique global optimal solution and is also a strong local minimum.

## Minimizing a Convex Function

### Theorem

Let $f : R^n \to R$ be a convex function, and let $S$ be a nonempty convex set in $R^n$. Consider the problem to minimize $f(\mathbf{x})$ subject to $\mathbf{x} \in S$. The point $\bar{\mathbf{x}} \in S$ is an optimal solution to problem if and only if it has a subgradient $\xi$ at $\bar{\mathbf{x}}$ such that $\xi(\mathbf{x} - \bar{\mathbf{x}}) \geq 0$ for all $\mathbf{x} \in S$.

## Minimizing a Convex Function

### Theorem

Consider the problem to minimize $f(\mathbf{x})$ subject to $\mathbf{x} \in S$, where $f$ is a convex and twice differentiable function and $S$ is a convex set, and suppose that there exists an optimal solution $\bar{\mathbf{x}}$. Then the set of alternative optimal solutions is characterized by the set

$$S^* = \{\mathbf{x} \in S : \nabla f(\bar{\mathbf{x}})^t(\mathbf{x} - \bar{\mathbf{x}}) \leq 0; \nabla f(\mathbf{x}) = \nabla f(\bar{\mathbf{x}})\}$$

## Maximizing a Convex Function

### Theorem

Let $f : R^n \to R$ be a convex function, and let $S$ be a nonempty convex set in $R^n$. Consider the problem to maximize $f(\mathbf{x})$ subject to $\mathbf{x} \in S$. If $\bar{\mathbf{x}} \in S$ is a local optimal solution to the problem, $\xi(\mathbf{x} - \bar{\mathbf{x}}) \leq 0$ for all $\mathbf{x} \in S$ where $\xi$ is any subgradient of $f$ at $\bar{\mathbf{x}}$.

## Maximizing a Convex Function

### Theorem

Let $f : R^n \to R$ be a convex function, and let $S$ be a nonempty compact polyhedral set in $R^n$. Consider the problem to maximize $f(\mathbf{x})$ subject to $\mathbf{x} \in S$. An optimal solution $\bar{\mathbf{x}}$ to the problem exists, where $\bar{\mathbf{x}}$ is an extreme point of $S$.

## Generalizations of a Convex Function

- Quasiconvex (Quasiconcave and Quasimonotone) Functions
- Differentiable Quasiconvex Functions
- Strictly Quasiconvex Functions
- Strongly Quasiconvex Functions
- Pseudoconvex Functions and Its Variants :-)
- Convexity at a Point

## Summary

- Definitions and Basic Properties
- Subgradients of Convex Functions
- Differentiable Convex Functions
- Minima and Maxima of Convex Functions
- Generalizations of Convex Functions

Thanks! Questions?