Convex Optimization and Congestion Control

Fabian Wirth

University of Würzburg

23.07. – 03.08.2012

Outline

• Part I:	Convexity and Convex Functions	Lectures 1, 2, 3
• Part II:	Convex Optimization	Lectures 4 and 5
• Part III:	Numerical Methods	Lectures 6 and 7
• Part IV:	Congestion Control	Lecture 8
• Part V:	Utility Based Congestion Control	Lecture 9
• Part VI:	Miscellaneous Problems in Networks	Lecture 10 (we shall see)

Part I: Convexity and Convex Functions

- I.1: Convex Sets
- I.2: Operations on Convex Sets and Construction of Convex Sets
- I.3: Separation
- I.4: Faces , Extreme Points and Recession Cones
- I.5: Duality
- I.6: Convex Functions
- I.7: Subgradients
- I.8: Optimality

Outline

Convex Functions

We now consider functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$.

Definition

The *domain* of such a function is defined as the set, where the function has finite values, i.e.

$$\mathrm{dom}\, f := \{x \in \mathbb{R}^n \mid f(x) \in \mathbb{R}\}.$$

Convex Functions

We now consider functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$.

Definition

The *domain* of such a function is defined as the set, where the function has finite values, i.e.

$$\operatorname{dom} f := \{x \in \mathbb{R}^n \mid f(x) \in \mathbb{R}\}.$$

With this notation we may as well consider the finite function

$$f : \operatorname{dom} f \subset \mathbb{R}^n \to \mathbb{R}.$$

Convex Functions

Definition

A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is called *convex*, if

(i) $\mathrm{dom}\, f$ is convex,

(ii) for all $x, y \in \mathrm{dom}\, f$, $\lambda \in [0, 1]$ we have

$$\lambda f(x) + (1 - \lambda) f(y) \geq f(\lambda x + (1 - \lambda) y) .$$

The function f is called *strictly convex* if in addition the inequality is strict, whenever $x \neq y$, $\lambda \in (0, 1)$.

Convex Functions

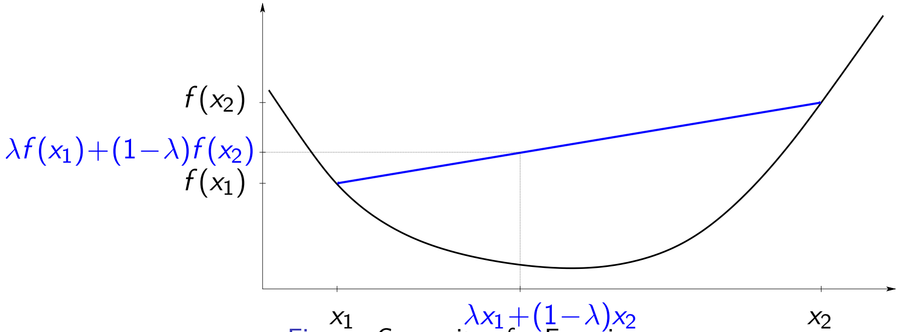

Figure: Convexity of a Function

The Epigraph

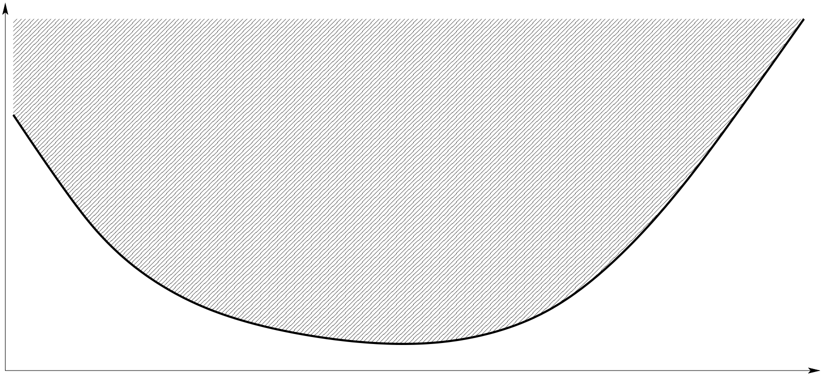

Figure: The epigraph of a convex function is convex.

Theorem

Consider $f : \mathrm{dom}\, f \subset \mathbb{R}^n \to \mathbb{R}$ and assume that $\mathrm{dom}\, f = \mathrm{ri}\,\mathrm{dom}\, f$ is convex.

(i) If f is continuously differentiable on $\mathrm{dom}\, f$, then f is convex if and only if

$$f(x) + \langle y - x, \nabla f(x) \rangle \leq f(y)\,, \qquad \text{for all } x, y \in \mathrm{dom}\, f\,. \tag{1}$$

If the inequality is strict whenever $x \neq y$, then f is strictly convex.

Convexity and Differentiability

Theorem

Consider $f : \operatorname{dom} f \subset \mathbb{R}^n \to \mathbb{R}$ and assume that $\operatorname{dom} f = \operatorname{ri} \operatorname{dom} f$ is convex.

(i) If f is continuously differentiable on $\operatorname{dom} f$, then f is convex if and only if

$$f(x) + \langle y - x, \nabla f(x) \rangle \leq f(y) \,, \qquad \text{for all } x, y \in \operatorname{dom} f \,. \tag{1}$$

If the inequality is strict whenever $x \neq y$, then f is strictly convex.

(ii) Assume f is twice continuously differentiable. Then f is convex if and only if the Hessian of f satisfies

$$Hf(x) \geq 0 \,, \qquad \text{for all } x \in \operatorname{dom} f \,. \tag{2}$$

If the Hessian is positive definite everywhere on $\operatorname{dom} f$, then f is strictly convex.

The Epigraph and Differentiability

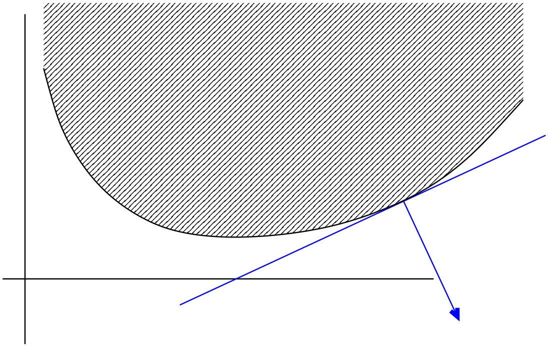

Figure: Supporting hyperplane of the epigraph.

Operations Preserving Convexity

Lemma

Let $f_1, f_2 : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex. Then

(i) $f_1 + f_2$ is convex.

(ii) $\max\{f_1, f_2\} : x \mapsto \max\{f_1(x), f_2(x)\}$ is convex.

Operations Preserving Convexity

Lemma

Let $f_1, f_2 : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex. Then

(i) $f_1 + f_2$ is convex.

(ii) $\max\{f_1, f_2\} : x \mapsto \max\{f_1(x), f_2(x)\}$ is convex.

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex, then for $x \in \mathrm{ri}\,\mathrm{dom}\, f$

$$f(x) = \max\{\langle x, c\rangle \mid c \in \mathbb{R}^n \text{ such that } \langle \cdot, c\rangle \leq f(\cdot)\}\,. \tag{3}$$

Note the right hand side is the maximization of linear (and therefore convex) functions. Thus the right hand side is automatically convex.

Continuity and Lipschitz Continuity

Definition

Let $U \subset \mathbb{R}^n$ be a nonempty set.
$f : U \to \mathbb{R}$ is called *locally Lipschitz continuous*, if for every $x \in U$ there is an $\varepsilon > 0$ and a constant L such that for all $y \in U \cap B_\varepsilon(x)$ it holds that

$$|f(x) - f(y)| \leq L\|x - y\| \,.$$

If the constant L can be chosen so that the inequality holds for all $x, y \in U$, then f is called globally Lipschitz continuous on U.

Further Estimates

Lemma

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and $x \in \operatorname{dom} f$. Then for any $y \in \mathbb{R}^n, 0 < m \leq l, 0 < h \leq k$ we have

$$\frac{f(x) - f(x - ly)}{l} \leq \frac{f(x) - f(x - my)}{m} \leq \frac{f(x + hy) - f(x)}{h} \leq \frac{f(x + ky) - f(x)}{k}.$$

Continuity and Lipschitz Continuity

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex. Then at every point $x \in \mathrm{ri}\, dom f$, the function f is continuous with respect to $\mathrm{aff}\,\mathrm{dom}\, f$.

Continuity and Lipschitz Continuity

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex. Then at every point $x \in \mathrm{ri}\, domf$, the function f is continuous with respect to $\mathrm{aff}\,\mathrm{dom}\, f$.

Theorem

If $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex, then f is locally Lipschitz continuous on $\mathrm{ri}\,\mathrm{dom}\, f$.

An Important Consequence

Rademacher's Theorem

Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$ be locally Lipschitz continuous. Then the set of points where f is not differentiable is a Lebesgue null set.

Lipschitz continuous functions are almost everywhere differentiable.

An Important Consequence

Rademacher's Theorem

Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$ be locally Lipschitz continuous. Then the set of points where f is not differentiable is a Lebesgue null set.

Lipschitz continuous functions are almost everywhere differentiable.

Corollary

A convex function f is almost everywhere differentiable on $\mathrm{ri}\,\mathrm{dom}\, f$.

Outline

Convexity and Differentiability

Remember:

Theorem

Consider $f : \operatorname{dom} f \subset \mathbb{R}^n \to \mathbb{R}$ and assume that $\operatorname{dom} f = \operatorname{ri} \operatorname{dom} f$ is convex.

(i) If f is continuously differentiable on $\operatorname{dom} f$, then f is convex if and only if

$$f(x) + \langle y - x, \nabla f(x) \rangle \leq f(y) \,, \qquad \text{for all } x, y \in \operatorname{dom} f \,. \tag{4}$$

If the inequality is strict whenever $x \neq y$, then f is strictly convex.

Remember:

Theorem

Consider $f : \operatorname{dom} f \subset \mathbb{R}^n \to \mathbb{R}$ and assume that $\operatorname{dom} f = \operatorname{ri} \operatorname{dom} f$ is convex.

(i) If f is continuously differentiable on $\operatorname{dom} f$, then f is convex if and only if

$$f(x) + \langle y - x, \nabla f(x) \rangle \leq f(y)\,, \qquad \text{for all } x, y \in \operatorname{dom} f\,. \tag{4}$$

If the inequality is strict whenever $x \neq y$, then f is strictly convex.

Supporting hyperplanes of the epigraph always exist!

Instead of assuming differentiability, we will now look for vectors p satisfying (4), when we replace $\nabla f(x)$ by p.

Subgradients

Definition

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex and $x \in \operatorname{dom} f$. A vector $p \in \mathbb{R}^n$ is called a subgradient vector of f in x, if $\operatorname{aff} \operatorname{dom} f$ recedes in direction p and

$$f(x) + \langle y - x, p \rangle \leq f(y)\,, \qquad \text{for all } y \in \operatorname{dom} f\,. \tag{5}$$

The subgradient of f at x is now a set defined by

$$\partial f(x) := \{ p \in \mathbb{R}^n \mid p \text{ satisfies } (5) \}\,. \tag{6}$$

Subgradients and Supporting Hyperplanes

The subgradients p are precisely the vectors for which $(p^\top, -1)^\top$ defines a supporting hyperplane of the epigraph.

Subgradients and Supporting Hyperplanes

The subgradients p are precisely the vectors for which $(p^\top, -1)^\top$ defines a supporting hyperplane of the epigraph.

The condition that $\operatorname{aff} \operatorname{dom} f$ recedes in direction p is only necessary, if $\operatorname{dom} f$ is a lower dimensional set. It is automatic, if $\operatorname{dom} f$ has interior points.

Subgradients and Supporting Hyperplanes

The subgradients p are precisely the vectors for which $(p^\top, -1)^\top$ defines a supporting hyperplane of the epigraph.

The condition that $\operatorname{aff}\operatorname{dom} f$ recedes in direction p is only necessary, if $\operatorname{dom} f$ is a lower dimensional set. It is automatic, if $\operatorname{dom} f$ has interior points.

Lemma

For every point in $\operatorname{ri}\operatorname{dom} f$ the subgradient of f is a nonempty set.

The Subgradient

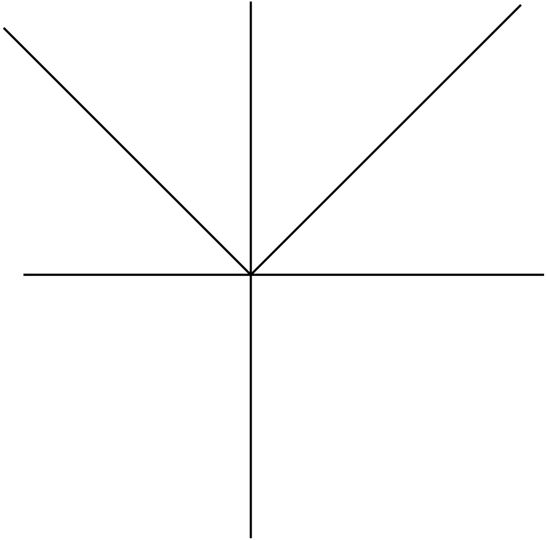

Figure: The absolute value.

The Subgradient

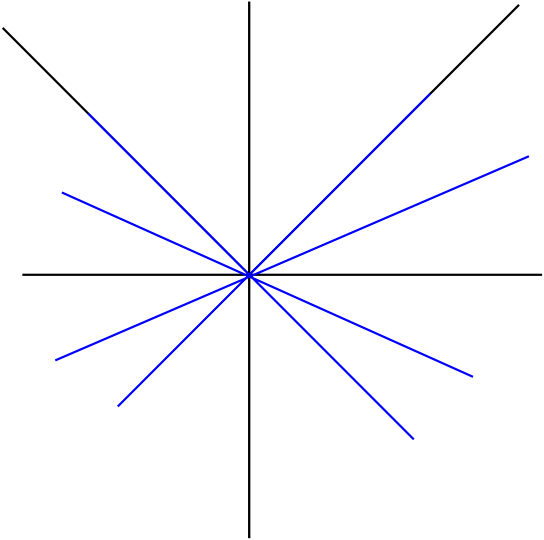

Figure: Supporting hyperplanes of the absolute value at $x = 0$.

Example: The Absolute Value

For the function $f : x \mapsto |x|$, we have differentiability at $x \neq 0$. The subgradient at zero is

$$\partial f(0) = [-1, 1]$$

because for $p \in [-1, 1]$ we have

$$f(0) + p \cdot (y - 0) = py \leq f(y) = |y| \, .$$

If $p \notin [-1, 1]$, then this inequality fails, for y with $py = |p||y|$.

Properties of the Subgradient

Lemma

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex. Then for all $x \in \mathrm{ri}\,\mathrm{dom}\, f$,

(i) the subgradient $\partial f(x)$ is bounded, closed, convex and nonempty.

(ii) if f is differentiable at x, then the subgradient contains the gradient of f at x as its only element, i.e.,

$$\partial f(x) = \{\nabla f(x)\}$$

The set-valued gradient map

Given $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ we now want to understand the behavior of the map

$$x \mapsto \partial f(x)\,.$$

The set-valued gradient map

Given $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ we now want to understand the behavior of the map

$$x \mapsto \partial f(x) .$$

Definition

Let $D \subset \mathbb{R}^n$ be a nonempty set and consider a set-valued map

$$F : D \to \{ \text{ nonempty, compact subsets of } \mathbb{R}^m \ \} .$$

The map F is called *upper semicontinuous* at $x \in D$, if for every sequence $x_k \to x$ in D and every convergent sequence $\{y_k\}_{k\in\mathbb{N}}$ in $\mathbb{R}^m$ satisfying

$$y_k \in F(x_k) , \qquad \text{for all } k,$$

we have $\lim_{k\to\infty} y_k \in F(x)$.
The set-valued map F is called upper semicontinuous, if it is upper semicontinuous at every point.

The set-valued gradient map

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex, then the set-valued map

$$x \mapsto \partial f(x)$$

is upper semicontinuous on $\mathrm{ri}\,\mathrm{dom}\, f$.

Gradients Say It All

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex and $x \in \mathrm{ri}\,\mathrm{dom}\, f$. Then

$$\partial f(x) = \mathrm{conv}\left\{ p \in \mathbb{R}^n \mid \exists x_k \in \mathrm{ri}\,\mathrm{dom}\, f, x_k \to x, p = \lim_{k\to\infty} \nabla f(x_k) \right\} .$$

In formulating the previous description we assume that $\nabla f(x_k)$ exists, if we write it down.

Subgradient Calculus

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be convex and $\lambda > 0$, then

$$\partial(\lambda f)(x) = \lambda \partial f(x)\,, \quad \forall x \in \operatorname{dom} f\,.$$

Theorem

Let $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $i = 1, \ldots, k$ be convex and define the convex function

$$F(x) := \max\{f_1(x), \ldots, f_k(x)\}\,.$$

For $x \in \cap_{i=1}^k \operatorname{ri} \operatorname{dom} f_i$ we have

$$\partial F(x) = \partial\left(\max\{f_1, \ldots, f_k\}\right)(x) = \operatorname{conv} \bigcup_{f_i(x)=F(x)} \partial f_i(x)\,.$$

The Summation Rule

Theorem (Moreau, Rockafellar)

If f, g are both convex and bounded on $D \subset \mathbb{R}^n$, D convex, then for all $x \in D$ we have

$$\partial(f+g)(x) = \partial f(x) + \partial g(x)\,, \tag{7}$$

where the sum on the right is the Minkowski sum of the (convex) subgradients of f and g at x.