Introduction to Optimization

Lecture 11: Convergence of the proximal-gradient algorithm. Conjugate functions.

university of groningen

Proximal-gradient algorithm

Suppose we want to find the minima of $f = g + h$, where $g : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is closed and convex, and $h : \mathbb{R}^N \to \mathbb{R}$ is L-smooth and convex.

Proximal-gradient algorithm

Suppose we want to find the minima of $f = g + h$, where $g : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is closed and convex, and $h : \mathbb{R}^N \to \mathbb{R}$ is L-smooth and convex.

Example

A typical example in image and signal processing, statistics, ML, is

$$f(x) = \frac{1}{2}\|Ax - b\|^2 + \rho\|x\|_1$$

for $x \in \mathbb{R}^N$.

Proximal-gradient algorithm

The proximal-gradient method consists in applying proximal iterations while linearizing the smooth function:

$$x_{n+1} = \operatorname{prox}_{\alpha g}\left(x_n - \alpha \nabla h(x_n)\right)$$

Proximal-gradient algorithm

The proximal-gradient method consists in applying proximal iterations while linearizing the smooth function:

$$
\begin{aligned}
x_{n+1} &= \operatorname{prox}_{\alpha g}\left(x_n - \alpha \nabla h(x_n)\right) \\
&= \operatorname{Argmin}\left\{ g(x) + \frac{1}{2\alpha}\left\| x - \left(x_n - \alpha \nabla h(x_n)\right)\right\|^2 \right\}.
\end{aligned}
$$

Proximal-gradient algorithm

The proximal-gradient method consists in applying proximal iterations while linearizing the smooth function:

$$
\begin{aligned}
x_{n+1} &= \operatorname{prox}_{\alpha g}\left(x_n - \alpha \nabla h(x_n)\right) \\
&= \operatorname{Argmin}\left\{ g(x) + \frac{1}{2\alpha}\left\| x - \left(x_n - \alpha \nabla h(x_n)\right)\right\|^2 \right\}.
\end{aligned}
$$

This subproblem has a unique solution characterized by

$$
0 \in \partial g(x_{n+1}) + \nabla h(x_n) + \frac{1}{\alpha}(x_{n+1} - x_n).
$$

Convergence of proximal-gradient sequences

Theorem

Let $f = g + h$, where $g : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is closed and convex, and $h : \mathbb{R}^N \to \mathbb{R}$ is L-smooth and convex. Take $\alpha \in (0, 1/L]$ and define (x_n) by

$$x_{n+1} = \mathrm{prox}_{\alpha g}\big(x_n - \alpha \nabla h(x_n)\big), \qquad n \geq 0.$$

If $S \neq \emptyset$, x_n converges to an $\hat{x} \in S$, and

$$f(x_n) - \min(f) \leq \frac{dist(x_0, S)^2}{2\alpha n}, \qquad n \geq 1.$$

Moreover, $\lim\limits_{n\to\infty} n\big(f(x_n) - \min(f)\big) = 0.$

Sketch of the proof

$f(x_n) + g(x_n)$ is nonincreasing

$$\begin{cases} f(x_{n+1}) & \leq & f(x_n) + \nabla f(x_n) \cdot (x_{n+1} - x_n) + \frac{L}{2}\|x_{n+1} - x_n\|^2 \\ g(x_{n+1}) & \leq & g(x_n) + \left(\frac{x_n - x_{n+1}}{\alpha} - \nabla f(x_n)\right) \cdot (x_{n+1} - x_n). \end{cases}$$

Sketch of the proof

$f(x_n) + g(x_n)$ is nonincreasing

$$\left\{\begin{array}{rcl} f(x_{n+1}) & \leq & f(x_n) + \nabla f(x_n)\cdot(x_{n+1}-x_n) + \frac{L}{2}\|x_{n+1}-x_n\|^2 \\ g(x_{n+1}) & \leq & g(x_n) + \left(\frac{x_n - x_{n+1}}{\alpha} - \nabla f(x_n)\right)\cdot(x_{n+1}-x_n). \end{array}\right.$$

Compatibility

$$\Big\|x_{n+1} - x_n - \alpha\big(\nabla f(x_{n+1}) - \nabla f(x_n)\big)\Big\|^2 \leq \|x_{n+1}-x_n\|^2.$$

Sketch of the proof

$f(x_n) + g(x_n)$ is nonincreasing

$$\begin{cases} f(x_{n+1}) & \leq & f(x_n) + \nabla f(x_n) \cdot (x_{n+1} - x_n) + \frac{L}{2}\|x_{n+1} - x_n\|^2 \\ g(x_{n+1}) & \leq & g(x_n) + \left(\frac{x_n - x_{n+1}}{\alpha} - \nabla f(x_n)\right) \cdot (x_{n+1} - x_n). \end{cases}$$

Compatibility

$$\left\| x_{n+1} - x_n - \alpha\big(\nabla f(x_{n+1}) - \nabla f(x_n)\big) \right\|^2 \leq \|x_{n+1} - x_n\|^2.$$

Convergence rate

$$\begin{cases} f(x_n) & \leq & f(p) + \nabla f(x_n) \cdot (x_n - p) \\ g(x_{n+1}) & \leq & g(p) + \left(\frac{x_n - x_{n+1}}{\alpha} - \nabla f(x_n)\right) \cdot (x_{n+1} - p). \end{cases}$$

Break

The Fenchel conjugate

The Fenchel conjugate of a closed convex function $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is the function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in \mathbb{R}^N} \{x^* \cdot x - f(x)\}.$$

The Fenchel conjugate

The Fenchel conjugate of a closed convex function $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is the function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in \mathbb{R}^N} \{x^* \cdot x - f(x)\}.$$

Examples

- In $\mathbb{R}$: (1) $f(x) = x$;

The Fenchel conjugate

The Fenchel conjugate of a closed convex function $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is the function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in \mathbb{R}^N} \{x^* \cdot x - f(x)\}.$$

Examples

- In $\mathbb{R}$: (1) $f(x) = x$; (2) $f(x) = \frac{1}{2}x^2$;

The Fenchel conjugate

The Fenchel conjugate of a closed convex function $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is the function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in \mathbb{R}^N} \{x^* \cdot x - f(x)\}.$$

Examples

- In $\mathbb{R}$: (1) $f(x) = x$; (2) $f(x) = \frac{1}{2}x^2$; (3) $f(x) = e^x$.

The Fenchel conjugate

The Fenchel conjugate of a closed convex function $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is the function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in \mathbb{R}^N} \{x^* \cdot x - f(x)\}.$$

Examples

- In $\mathbb{R}$: (1) $f(x) = x$; (2) $f(x) = \frac{1}{2}x^2$; (3) $f(x) = e^x$.
- In $\mathbb{R}^N$: (4) $f(x) = c \cdot x + \alpha$;

The Fenchel conjugate

The Fenchel conjugate of a closed convex function $f : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ is the function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(x^*) = \sup_{x \in \mathbb{R}^N} \{x^* \cdot x - f(x)\}.$$

Examples

- In $\mathbb{R}$: (1) $f(x) = x$; (2) $f(x) = \frac{1}{2}x^2$; (3) $f(x) = e^x$.
- In $\mathbb{R}^N$: (4) $f(x) = c \cdot x + \alpha$; (5) $f(x) = \phi(\|x\|)$.

Basic properties

Proposition

1. *The function f^* is closed, convex and not identically $+\infty$.*

Basic properties

Proposition

1. *The function* f^* *is closed, convex and not identically* $+\infty$.
2. *If* $f \leq g$, *then* $f^* \geq g^*$.

Basic properties

Proposition

1. *The function f^* is closed, convex and not identically $+\infty$.*
2. *If $f \leq g$, then $f^* \geq g^*$.*
3. *Fenchel-Young Inequality:* $f(x) + f^*(x^*) \geq \langle x^*, x \rangle$.

Basic properties

Proposition

1. *The function f^* is closed, convex and not identically $+\infty$.*
2. *If $f \leq g$, then $f^* \geq g^*$.*
3. *Fenchel-Young Inequality: $f(x) + f^*(x^*) \geq \langle x^*, x \rangle$.*
   There is equality if, and only if, $x^* \in \partial f(x)$.

Basic properties

Proposition

1. *The function* f^* *is closed, convex and not identically* $+\infty$.
2. *If* $f \leq g$, *then* $f^* \geq g^*$.
3. *Fenchel-Young Inequality*: $f(x) + f^*(x^*) \geq \langle x^*, x \rangle$.
   There is equality if, and only if, $x^* \in \partial f(x)$.
4. $f^{**} = f$.

Basic properties

Proposition

1. *The function* f^* *is closed, convex and not identically* $+\infty$.
2. *If* $f \leq g$, *then* $f^* \geq g^*$.
3. *Fenchel-Young Inequality:* $f(x) + f^*(x^*) \geq \langle x^*, x\rangle$.
   There is equality if, and only if, $x^* \in \partial f(x)$.
4. $f^{**} = f$.
5. *Legendre-Fenchel Reciprocity Formula:* $x^* \in \partial f(x)$ *if, and only if,* $x \in \partial f^*(x^*)$.

Basic properties

Proposition

1. *The function f^* is closed, convex and not identically $+\infty$.*
2. *If $f \leq g$, then $f^* \geq g^*$.*
3. *Fenchel-Young Inequality: $f(x) + f^*(x^*) \geq \langle x^*, x \rangle$. There is equality if, and only if, $x^* \in \partial f(x)$.*
4. $f^{**} = f$.
5. *Legendre-Fenchel Reciprocity Formula: $x^* \in \partial f(x)$ if, and only if, $x \in \partial f^*(x^*)$.*
6. *Let $\mu\ell = 1$. Then, f is μ-strongly convex if, and only if, f^* is ℓ-smooth.*