Supplemental Material for 10/30 Book Reading Seminar

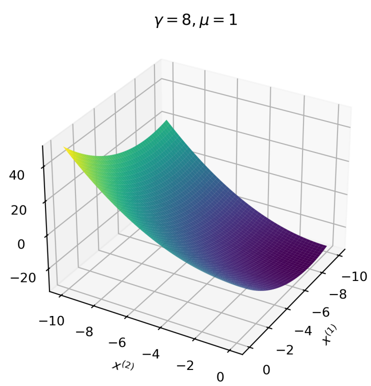
Hiroki Hamaguchi

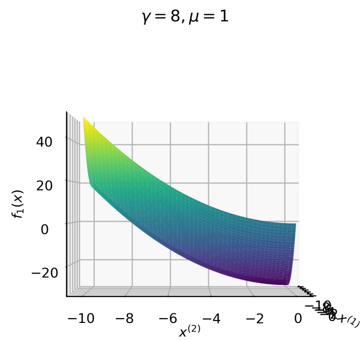
October 29, 2024

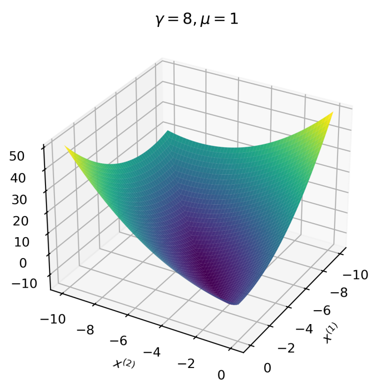
Abstract

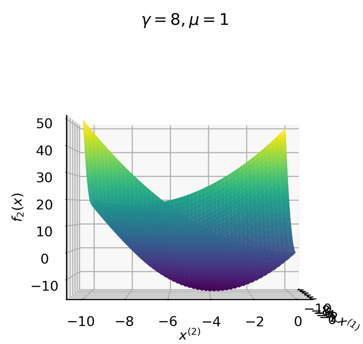
This is the supplemental material for the book reading seminar. Please refer to the textbook and the whiteboard for the main content. If necessary, please also refer to the repository.

3.2.1







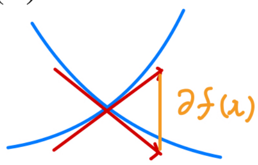


Definition 3.1.5 A vector g is called a *subgradient* of the function f at the point $x_0 \in \operatorname{dom} f$ if for any $y \in \operatorname{dom} f$ we have

$$f(y) \geq f(x_0) + \langle g, y - x_0 \rangle. \tag{3.1.23}$$

The set of all subgradients of f at x_0, $\partial f(x_0)$, is called the *subdifferential* of the function f at the point x_0.

Lemma 3.1.13 *Let functions f_i, $i = 1 \ldots m$, be closed and convex. Then the function $f(x) = \max\limits_{1 \leq i \leq m} f_i(x)$ is closed and convex. For any $x \in \operatorname{int}(\operatorname{dom} f) = \bigcap\limits_{i=1}^{m} \operatorname{int}(\operatorname{dom} f_i)$, we have*



$$\partial f(x) = \operatorname{Conv}\{\partial f_i(x) \mid i \in I(x)\}, \tag{3.1.36}$$

where $I(x) = \{i : f_i(x) = f(x)\}$.

Figure 1: p.195: "described in Sect. 3.1.6"

Definition 3.1.6 Let the set $X \subseteq \operatorname{dom} f$ be closed and convex. The set

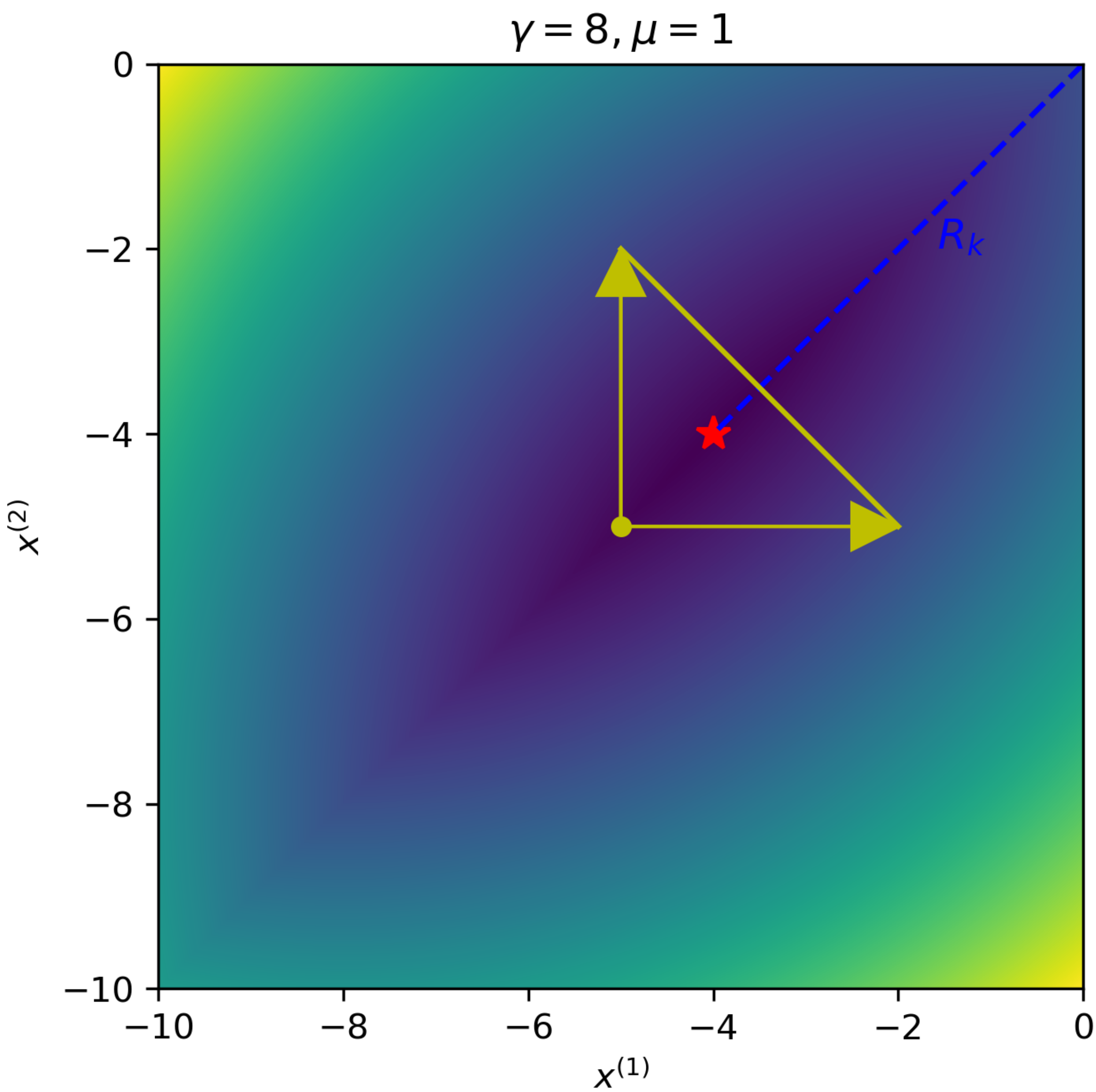
$$\widehat{\partial f}(X) = \bigcap_{x \in X} \partial f(x) \tag{3.1.32}$$

is called the *epigraph facet* of the set X.

Theorem 3.1.20 *Let $X^* = Arg \min\limits_{x \in \operatorname{dom} f} f(x)$. Then a closed convex set X_* is a subset of X^* if and only if*

$$0 \in \widehat{\partial f}(X_*).$$

Figure 2: p.196: "Further, by Theorem 3.1.20,"



3.2.2

Corollary 3.1.6 *Let $Q \subseteq \operatorname{dom} f$ be a closed convex set, $x_0 \in Q$, and*

$$x^* \in \operatorname*{Arg\,min}_{x \in Q} f(x).$$

Then for any $g \in \partial f(x_0)$, we have $\langle g, x_0 - x^* \rangle \geq 0$. $\square$

Figure 3: p.199: "We have justified this property in Corollary 3.1.6."

3.2.3

Definition 2.2.2 Let Q be a closed set and $x_0 \in \mathbb{R}^n$. Define

$$\pi_Q(x_0) = \arg\min_{x \in Q} \parallel x - x_0 \parallel . \quad (2.2.46)$$

We call $\pi_Q(x_0)$ the *Euclidean projection* of the point x_0 onto the set Q.

Figure 4: p.202: "π_Q"

Lemma 2.2.8 *For any two point $x \in Q$ and $y \in \mathbb{R}^n$, we have*

$$\parallel x - \pi_Q(y) \parallel^2 + \parallel \pi_Q(y) - y \parallel^2 \leq \parallel x - y \parallel^2 . \quad (2.2.49)$$

Proof Indeed, in view of (2.2.47), we have

$$\parallel x - \pi_Q(y) \parallel^2 - \parallel x - y \parallel^2 = \langle y - \pi_Q(y), 2x - \pi_Q(y) - y \rangle$$

$$\leq - \parallel y - \pi_Q(y) \parallel^2 .$$

□

Figure 5: p.202: "Then, in view of Lemma 2.2.8,"

Theorem 3.2.1 *For any class $\mathscr{P}(x_0, R, M)$ and any k, $0 \leq k \leq n-1$, there exists a function $f \in \mathscr{P}(x_0, R, M)$ such that*

$$f(x_k) - f^* \geq \frac{MR}{2(2+\sqrt{k+1})}$$

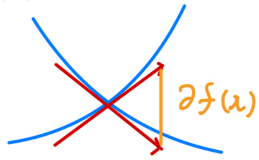
for any optimization scheme, which generates a sequence $\{x_k\}$ satisfying the condition

$$x_k \in x_0 + \mathrm{Lin}\{g(x_0), \ldots, g(x_{k-1})\}.$$

Figure 6: p.204: "with the lower bound of Theorem 3.2.1"

3.2.4

Lemma 3.1.13 *Let functions* f_i, $i = 1 \dots m$, *be closed and convex. Then the function* $f(x) = \max_{1 \le i \le m} f_i(x)$ *is closed and convex. For any* $x \in \operatorname{int}(\operatorname{dom} f) = \bigcap_{i=1}^{m} \operatorname{int}(\operatorname{dom} f_i)$, *we have*



$$\partial f(x) = \operatorname{Conv}\{\partial f_i(x) \mid i \in I(x)\}, \tag{3.1.36}$$

where $I(x) = \{i : f_i(x) = f(x)\}$.

Figure 7: p.205: “we can do so for the functions f_j (see Lemma 3.1.13)”

Lemma 2.2.8 *For any two point* $x \in Q$ *and* $y \in \mathbb{R}^n$, *we have*

$$\| x - \pi_Q(y) \|^2 + \| \pi_Q(y) - y \|^2 \le \| x - y \|^2 . \tag{2.2.49}$$

Proof Indeed, in view of (2.2.47), we have

$$\| x - \pi_Q(y) \|^2 - \| x - y \|^2 = \langle y - \pi_Q(y), 2x - \pi_Q(y) - y \rangle$$

$$\le - \| y - \pi_Q(y) \|^2 .$$

□

Figure 8: p.206: “(2.2.49)”

Theorem 3.2.1 *For any class* $\mathscr{P}(x_0, R, M)$ *and any* k, $0 \le k \le n-1$, *there exists a function* $f \in \mathscr{P}(x_0, R, M)$ *such that*

$$f(x_k) - f^* \ge \frac{MR}{2(2+\sqrt{k+1})}$$

for any optimization scheme, which generates a sequence $\{x_k\}$ *satisfying the condition*

$$x_k \in x_0 + \operatorname{Lin}\{g(x_0), \dots, g(x_{k-1})\}.$$

Figure 9: p.206: “with the result of Theorem 3.2.1”