1 Algorithm Descriptions and Proofs

Consider the optimization problem $\min_{x\in C} f(x)$ where the feasible set $C = \{x \in \mathbb{R}^n \mid g(x) \geq 0\}$, the objective function $f : \mathbb{R}^n \to \mathbb{R}$, and the constraints are described by the function $g : \mathbb{R}^n \to \mathbb{R}^d$.
Assumption 1. The functions f, g are continuously differentiable and have a Lipschitz continuous gradient. $f(x) \to \infty$ when $|x| \to \infty$ and C is non-empty and bounded.
Assumption 2. The Mangasarian-Fromovitz constraint qualification (MFCQ) is satisfied for all $x \in \mathbb{R}^n$, i.e. $\forall x \in \mathbb{R}^d$, $\exists w \in \mathbb{R}^n$ s.t. $\nabla g_i(x)w > 0$ for all $i \in I_x$, where $I(x) = \{i \in \mathbb{Z} \mid g_i(x) \leq 0\}$.
Fact. The MFCQ condition is automatically satisfied for $x \in C$.

Tangent cones have a natural role in the theory of flow-invariant sets and gradient inclusions.

Definition 1. The Clarke's tangent cone of C contains all $\delta x \in T_C(x)$ if there exists two sequences $x_j \to x, x_j \in C, t_j \downarrow 0$ such that $(x_j - x)/t_j \to \delta x$. The normal cone is defined as follows: $N_C(x) = \{\lambda \in \mathbb{R}^n \mid \lambda^\top \delta x \leq 0, \forall \delta x \in T_C(x)\}$.

Lemma 1. Suppose that $x \in C$, then every $\delta x \in T_C(x)$ satisfies $\nabla g_i(x)\delta x \geq 0, \forall i \in I_x$. The converse also holds.

Proof. $(\Rightarrow)$: $\delta x \in T_C(x)$ implies that there exists two sequences $\{x_j\} \to x, \{x_j\} \subset C, t_j \downarrow 0$ for all $j \in \mathbb{N}$ and

$$\frac{x_j - x}{t_j} \to \delta x,$$

which implies that

$$\frac{g(x_j) - g(x)}{x_j - x} \cdot \frac{x_j - x}{t_j} \geq 0.$$

This is because $x_j \in C$ implies that $g_i(x_j) \geq 0$ and $g_i(x) \leq 0$ for all $i \in I(x)$.

$(\Leftarrow)$: Adapted from R. Herzog, 2023, a simplified version. Let δx satisfy $\nabla g_i(x)\delta x \geq 0, \forall i \in I_x$, also let δy be given by MFCQ such that $\nabla g_i(w)\delta y > 0, \forall i \in I(x)$. Put $\ell(t) := \delta x + t \cdot \delta y$. Then for all $t > 0$, we have $\nabla g_i(x)\ell(t) > 0, \forall i \in I(x)$, implying that $\ell(t)$ are all feasible MFCQ vectors.

Now, we claim that $\ell(t) \in T_C(x)$ for all $t \in \mathbb{R}_{++}$. Let $\gamma(t) := x + t\ell(t)$, $t \in (-\varepsilon, \varepsilon)$, for an infinitestimally small ε, given by the continuity of g. Then, $y(t) \in C$ for every $t \in [0, \varepsilon)$ and $\gamma(0) = x, \gamma'(0) = \ell(t)$. For an arbitrary sequence $\{t_j\} \downarrow 0$ and $x_k = \gamma(t_j) \to x$ we have

$$\ell(t) = \gamma'(0) = \lim_{j\to\infty} \frac{\gamma(t_j) - \gamma(0)}{t_j - 0} = \lim_{j\to\infty} \frac{x_j - x}{t_j} \in T_C(x).$$

Since $T_C(x)$ is closed, $\delta x = \lim_{t\to 0} \ell(t) \in T(x)$. □

Now, we can simplify the tangent cone and the normal cone for any $x \in C$ as follows, due to the Mangasarian-Fromovitz constraint qualification:

$$T_C(x) = \{x \mid \nabla g_i(x)^\top x \geq 0, \forall i \in I_x\}, N_C(x) = \left\{ \lambda \in \mathbb{R}^d_+ \middle| - \sum_{i\in I(x)} \lambda_i \nabla g_i(x) \right\}.$$

Definition 2. Further define the set $V_\alpha(x) := \{v \in \mathbb{R}^n \mid \nabla g_i(x)^\top v + \alpha g_i(x) \geq 0, \forall i \in I_x\}$, where $\alpha > 0$. $V_\alpha(x)$ is guaranteed to be non-empty for any x.

Indeed. For the case where $x \in C$, $V_\alpha(x)$ is nothing but $T_C(x)$. Otherwise, consider any $g_i(x) < 0$, by MFCQ there exists u such that $\nabla g_i(x)^\top u > 0$ and therefore by scaling we have $\nabla g_i(x)^\top v \geq -\alpha g_i(x) \geq 0$.

Definition 3. The indicator function for a set C is defined as:

$$\psi_C(x) = \begin{cases} 0, & x \in C, \\ \infty, & \text{otherwise.} \end{cases}$$

Theorem 1. Let $x : [0, \infty) \to \mathbb{R}^n$ be an absolutely continuous trajectory with a piecewise continuous derivative. Then, for any $x(0) \in C$, the following are equivalent:

$$\begin{aligned}
&\dot{x}(t) := -\nabla f(x(t)) + R(t), -R(t) \in N_C(x(t)), && \forall t \in [0, \infty) \text{ almost everywhere}, \\
&\dot{x}(t)^+ := -\nabla f(x(t)) + R(t), -R(t) \in \partial \psi_{V_\alpha(x(t))}(\dot{x}(t)^+), && \forall t \in [0, \infty), \\
&\dot{x}(t)^+ := -\operatorname*{argmin}_{v \in V_\alpha(x(t))} \frac{1}{2} |v + \nabla f(x(t))|^2, && \forall t \in [0, \infty).
\end{aligned}$$

Lemma 2. Using the `ASkewSGD` algorithm with step sizes $\{\gamma_k\}$ of $\sum_{i=1}^{\infty} \gamma_i = \infty$, $\sum_{i=1}^{\infty} \gamma_i^2 < \infty$, the iterate $\{w_k\}$ is guaranteed to converge and $\lim_{k \to \inf} d(w_k, C_\varepsilon) = 0$.

Proof. See Leconte et al., 2023, Appendix A.3. $\square$

Lemma 3. Let $k_0 = \sup_{1 \leq i \leq d, 1 \leq j \leq K_i} \sup\{k : \gamma_k M \geq \max(c_- - \frac{c_j^i + c_{j+1}^i}{2}, -c_+ + \frac{c_j^i + c_{j+1}^i}{2})\}$. Since w must

2 Piecewise Convexity

The piecewise convexity can be granted given the changes applied to the constraints.

Definition 4. A function $F : \mathbb{R}^n \mapsto \mathbb{R}$ is called a piecewise convex function on $\mathbb{R}^n$ if it can be decomposed into:

$$F(x) = \min\{f_1(x), f_2(x), \ldots, f_m(x)\}$$

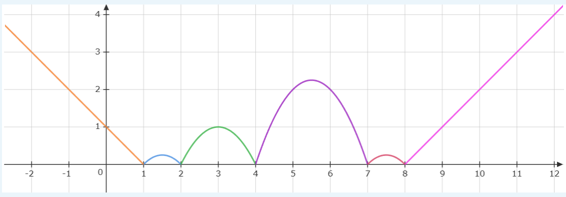
where $f_j : \mathbb{R}^n \mapsto \mathbb{R}$ are convex functions for all $j \in M = 1, 2, \ldots, m$.

Lemma 4. Given a set of K quantization levels $\mathcal{Q} = \{q_1, q_2, \ldots, q_K\}$. Define the piecewise function

$$\psi(w) := \begin{cases} (q_1 - w), & w < q_1, \\ (q_{i-1} - w)(w - q_i), & q_{i-1} \leq w < q_i, i = 2, \ldots, K \\ w - q_K, & w \geq q_K, \end{cases}$$

for all $w \in \mathbb{R}$. Then, $-\psi'$ is a piecewise convex function.

Proof. That is, to prove that ψ can be decomposed into concave functions $f_1, \ldots, f_m$ where $F(x) = \max(f_1(x), f_2(x), \ldots, f_m(x))$.



The proof is pictorially shown by the plot above. Select $m = K + 1$ such that the f_i's ($i = 1, \ldots, K + 1$) corresponds to the analytic continuation of every piece of function.

$$\begin{cases} f_1(w) := (q_1 - w), \\ f_i(w) := (q_{i-1} - w)(w - q_i), i = 2, \ldots, K \\ f_{K+1}(w) := w - q_K. \end{cases}$$

□

Note. We have changed the definition of ψ, the preliminary verification through proofs shows no problem of convergence failure. Apparently the new setup still satisfies the MFCQ condition for every $w \neq (q_i + q_{i+1})/2, \forall i = 1, \ldots, K - 1$. Also observed that this change does not alter the convergence property for the logistic regression problem. Now, we should note that

$$0 < \varepsilon \leq \inf_{1 \leq i \leq d} \inf_{1 \leq j \leq K^i} \left| c_j^i - c_{j+1}^i \right|^2 / 4$$

ensures the disconnectedness of the set C_ε.

3 Lagrange Duality

The problem $\mathcal{P}$ of

$$\min_{x \in C} f(x), C = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$$

is equivalent to the primal problem

$$\inf_{x \in \mathbb{R}^n} \sup_{\lambda \succeq 0} f(x) + \sum_{i=1}^{d} \lambda_i g_i(x).$$

We consider the dual problem

$$\sup_{\lambda \succeq 0} \inf_{x \in \mathbb{R}^n} f(x) + \sum_{i=1}^{d} \lambda_i g_i(x).$$

Theorem 2. Suppose that x^* is a local minimizer of $\mathcal{P}$ which satisfies the MFCQ. Then there exist Lagrange multipliers λ^* (not necessary unique) such that the KKT conditions are satisfied. The set of Lagrange multipliers $\Lambda(x^*)$ is compact.

Therefore, the KKT points can be captured by the following set:

$$\mathcal{Z}_\varepsilon = \{w \in C_\varepsilon : 0 \in -\nabla \ell(w) + N_{C_\varepsilon}(w)\}$$

Theorem 3. If $f : \mathbb{R}^d \mapsto \mathbb{R}$ is twice continuously differentiable and satisfies the strict saddle property, then gradient descent with a random initialization and sufficiently small constant step size converges to a local minimizer or negative infinity almost surely. Call x a critical point of f if $\nabla f(x) = 0$, and say that f satisfies the strict saddle property if each critical point x of f is either a local minimizer, or a "strict saddle", i.e, $\nabla^2 f(x)$ has at least one strictly negative eigenvalue.(J. D. Lee, in PMLT, 2016)

Algorithm 1 Dual gradient ascent method (convex constraints)

1: Start with an initial dual guess $\lambda(0) \geq 0$.
2: **for** $k = 1, 2, \ldots$ **do**
3: $\quad x^{(k)} \in \operatorname*{argmin}_x \ell(x) + (\lambda^{(k-1)})^\top g(x)$
4: $\quad \lambda^{(k)} = \max\{\lambda^{(k-1)} + \gamma_k g(x^k), 0\}$
5: **end for**

Assumption 3. The objective function f is convex and continuously differentiable.

How we can find x^* efficiently, as $\nabla \ell$ is implicit?

The problem is that $f(x) - \lambda g$ is a combination of a convex and a concave function (where g is convex and $g \geq 0$ is required).

`https://proceedings.neurips.cc/paper_files/paper/2023/file/a961dea42c23c3c0d01b79918701fb6e-Paper pdf`

4 Discussions

4.1 More ideas Under the assumption of non-summable and square-summable step sizes, $\limsup_{k\to\infty} d(w_k, C_\varepsilon) = 0$ almost surely. Given that the current piece C_ε is convex, disconnected, can we guarantee that $d(w_k, \mathcal{Z}_\varepsilon)$ converges almost surely? Note that the direction picked by speed never guide w to leave the feasible set if any constraint is already violated. (Will continue to read Boob, 2019. *arXiv*: `https://arxiv.org/pdf/1908.02734`)

First, we want to know if the algorithm escapes from saddle points (so not just stationary points but minimizers). (Stochastic case follows from `https://sites.math.washington.edu/~ddrusv/aiming_deep.pdf` and `https://hal.science/hal-03442137/file/tame.pdf`, step sizes `https://hal.science/hal-02564349/file/clarke.pdf`)

Second, yet another Gradient Flow inspection. (`https://openreview.net/pdf?id=xuw7R0hP7G`)

Third, experiment on C1-smooth and non C2-smooth functions (don't really know if they are even usable). Example: finding out the cases when the algorithm fails to converge?

Fourth, another algorithm for non-convex optimization. (`https://arxiv.org/pdf/1908.02734`)

4.2 A survey of quantization methods

4.3 Updates on the numerical experiment

4.4 Incoming Events