Coursework (1) for *Introductory Lectures on Optimization*

Hanxuan Li
3220106039

Oct. 16, 2023

Excercise 1. Please provide several main optimization forms in reinforcement learning.

Solution of Excercise 1:

1. Linear optimization
   (a) objective: minimize or maximize a linear objective function subject to linear equality and inequality constraints.
   (b) formulation:
   $$Minimize\ c^T x\ subject\ to\ Ax \leq b,\ x \geq 0 \tag{1}$$
2. Quadratic optimization
   (a) objective: minimize or maximize a quadratic objective function subject to linear equality and inequality constraints.
   (b) formulation:
   $$Minimize\ x^T Qx + c^T x\ subject\ to\ Ax \leq b,\ x \geq 0 \tag{2}$$
3. Deep Q-Network(DQN)
   (a) objective: optimize the action-value function $Q(s, a; \theta)$
   (b) formulation:
   $$\underset{\theta}{min}\, \mathbb{E}_{s,a,r,s'} [(Q(s, a; \theta) - (r + \lambda \underset{a'}{max}\, Q(s^{'}, a^{'}; \theta^{-})))^2],\ with\ no\ explicit\ constraints \tag{3}$$
   Here, θ represents the network parameters, s is a state, a is an action, r is a reward, $s^{'}$ is the next state, λ is the discount factor, and θ^{-} are the parameters of the target network.

□

Excercise 2. Please investigate what other commonly used oracles exist in the field of optimization, besides zeroth-order, first-order, and second-order oracles.

Solution of Excercise 2:

1. Stochastic Gradient Oracles: instead of providing the exact gradient, these orcales provide an estimate of the gradient based on a subset of the data(mini-batch) or a noisy estimate of the objective function.

2. Stochastic Hessian Oracles: return stochastic or approximated Hessians.
3. Subdifferential Oracles: return subgradients(generalizations of gradients for non-smooth functions),which are used in non-smooth optimization.
4. Bayesian Oracles: utilize Bayesian optimization techniques that model the objective function as a probabilistic process.
5. Quasi-Newton Oracles: return approximations of the Hessianmatrix using rank-one updates.

□

Excercise 3. For the performance analysis of the Uniform Grid Method, Proof that

$$\left(\left\lfloor\frac{L}{2\epsilon}\right\rfloor+2\right)^n, \text{ and } \left(\left\lfloor\frac{L}{2\epsilon}\right\rfloor\right)^n,$$

coincide up to an absolute constant multiplicative factor if $\epsilon \leq O(\frac{L}{n})$.

Proof of Excercise 3: Since $\epsilon \leq O(\frac{L}{n})$,Then we have:

$$\frac{\left(\left\lfloor\frac{L}{2\epsilon}\right\rfloor+2\right)^n}{\left(\left\lfloor\frac{L}{2\epsilon}\right\rfloor\right)^n} = (1+\frac{1}{\frac{1}{2}\left\lfloor\frac{L}{2\epsilon}\right\rfloor})^n \leq (1+\frac{1}{cn})^n = ((1+\frac{1}{cn})^{cn})^{\frac{1}{c}} \underset{n\to\infty}{=} (e)^{\frac{1}{c}}, \tag{4}$$

where c is a constant, and we obtain the last equality from the following theorem:

$$\underset{n\to\infty}{lim} = (1+\frac{1}{n})^n = e \tag{5}$$

Thus we reach the conclusion that $\left(\left\lfloor\frac{L}{2\epsilon}\right\rfloor+2\right)^n$ and $\left(\left\lfloor\frac{L}{2\epsilon}\right\rfloor\right)^n$ coincide up to an absolute constant multiplicative factor if $\epsilon \leq O(\frac{L}{n})$. □