1 Outline

In this lecture, we cover

- continuous submodular function maximization,
- max cut problem.

2 Continuous Submodular Functions

In this section, we consider a continuous extension of the discrete submodular function maximization problem. It is not difficult to observe that the multilinear extension of a submodular set function satisfies the following **diminishing returns (DR) property**:

$$F(x+\delta e_i) - F(x) \geq F(y+\delta e_i) - F(y)$$

for any $i \in E$, $\delta \geq 0$, and $x, y \in [0,1]^{|E|}$ with $x \leq y$. The multilinear extension is neither convex nor concave. However, it is **concave along any nonnegative direction (up-concave)** $v \geq 0$, i.e.,

$$g(t) = F(x + t \cdot v)$$

is concave with respect to t. Moreover, the multilinear extension is **smooth**, i.e. there exists $\beta > 0$ such that

$$\|\nabla F(x) - \nabla F(y)\|_2 \leq \beta \|x - y\|_2.$$

Based on these properties, we extend submodularity to continuous functions.

We say that a function $F : \mathbb{R}^d \to \mathbb{R}$ is **(continuous) DR-submodular** if it satisfies the diminishing returns (DR) property. For some domain C,

$$F(x+\delta e_i) - F(x) \geq F(y+\delta e_i) - F(y)$$

for any $i \in [d]$, $\delta \geq 0$, and $x, y \in C$ with $x \leq y$.

Lemma 19.1 ([BMBK17]). *If F is DR-submodular, then it is up-concave.*

In addition, DR-submodular functions satisfy the following properties.

- A differentiable function is DR-submodular if and only if

$$\nabla F(x) \geq \nabla F(y) \quad \text{for any } x, y \text{ with } x \leq y.$$

- A twice-differentiable function is DR-submodular if and only if

$$\nabla^2 F(x) = \left(\frac{\partial^2 F(x)}{\partial x_i \partial x_j}\right) \leq 0.$$

On top of the multilinear extension of a submodular set function, there exist other examples of DR-submodular functions.

- Quadratic functions $x^\top Ax/2 + b^\top x + c$ with $A \leq 0$.
- $\sum_{i,j} \varphi_{i,j}(x_i - x_j)$ where $\varphi_{i,j}$ is convex for every i, j.
- $g(\sum_i w_i x_i)$ where g is concave and $w \geq 0$.
- $\text{logdet}(\sum_i x_i A_i)$ where each A_i is positive definite and $x \geq 0$.

Continuous DR-submodular arise in isotonic regression [Bac18], robust budget allocation [SJ17, SKIK14], online resource allocation [EF16], and adwords for e-commerce and advertising [DJ12, MSVV05]. Then we consider the following continuous DR-submodular maximization problem.

$$\text{maximize} \quad F(x) \quad \text{subject to} \quad x \in C$$

where F is continuous DR-submodular, F is monotone, i.e. $F(x) \leq F(y)$ for any x, y with $x \leq y$, and C is a convex constraint set. We further assume that $0 \in C$, C is bounded and **down-closed**, i.e., if $y \in C$, then any $0 \leq x \leq y$ belongs to C. if To solve the problem, we may the continuous greedy algorithm. We present the **conditional gradient method** due to Bian et al. [BMBK17], which is also referred to as the Frank-Wolfe algorithm for submodular maximization. Note that for

Algorithm 1 Conditional gradient method

Start with $x_0 = 0$.
for $t = 1, \ldots, T$ **do**
  Obtain $v_t \in \arg\max_{v \in C} \left\{\nabla F(x_{t-1})^\top v\right\}$.
  Update $x_t = x_{t-1} + (1/T)v_t$.
end for
Output x_T.

each t, we have

$$x_t = \frac{1}{T}\sum_{i=1}^{t} v_i = \frac{T-t}{T} \cdot 0 + \frac{1}{T}\sum_{i=1}^{t} v_i \in C$$

Theorem 19.2 ([BMBK17]). *Assume that F is monotone, β-smooth in the ℓ_2-norm, and DR-submodular. We further assume that $0 \in C$, C is down-closed, and $\|v\|_2 \leq R$ for any $v \in C$. Then x_T returned by conditional gradient (Algorithm 1) satisfies*

$$F(x_T) \geq \left(1 - \frac{1}{e}\right) \max_{x \in C} F(x) - \frac{\beta R^2}{2T}$$

under $F(0) = 0$.

Proof. Since F is β-smooth, we have

$$\begin{aligned} F(x_t) &\geq F(x_{t-1}) + \nabla F(x_{t-1})^\top (x_t - x_{t-1}) - \frac{\beta}{2}\|x_t - x_{t-1}\|_2^2 \\ &= F(x_{t-1}) + \frac{1}{T}\nabla F(x_{t-1})^\top v_t - \frac{\beta}{2T^2}\|v_t\|_2^2. \end{aligned} \tag{19.1}$$

Let x^* be an optimal solution, and let v^* be defined as

$$v^* := (x^* \vee x) - x = (x^* - x) \vee 0 \geq 0$$

where $p \vee q$ for two vectors p and q takes the coordinate-wise maximum values of p and q. Then we have $0 \leq v^* \leq x^*$, and the down-closedness of C implies that $v^* \in C$. Moreover, it follows from the monotonicity of F that

$$F(x + v^*) = F(x^* \vee x) \geq F(x^*). \tag{19.2}$$

Note that

$$\nabla F(x_{t-1})^\top v_t \geq \nabla F(x_{t-1})^\top v^* \geq F(x_{t-1} + v^*) - F(x_{t-1}) \geq F(x^*) - F(x_{t-1})$$

where the first inequality is due to our choice of v_T, the second inequality holds because $v^* \geq 0$ and F is up-concave, and the third one comes from (19.2). Combined with (19.1), we obtain

$$F(x_t) \geq \left(1 - \frac{1}{T}\right) F(x_{t-1}) + \frac{1}{T} F(x^*) - \frac{\beta}{2T^2} R^2.$$

This implies that

$$F(x^*) - F(x_t) \leq \left(1 - \frac{1}{T}\right) (F(x^*) - F(x_{t-1})) + \frac{\beta R^2}{2T^2}$$

Furthermore, it follows that

$$\begin{aligned} F(x^*) - F(x_T) &\leq \left(1 - \frac{1}{T}\right)^T (F(x^*) - F(x_0)) + \frac{\beta R^2}{2T^2} \sum_{t=1}^{T} \left(1 - \frac{1}{T}\right)^{t-1} \\ &\leq \frac{1}{e} (F(x^*) - F(x_0)) + \frac{\beta R^2}{2T}, \end{aligned}$$

as required. □

3 Max Cut Problem

We consider the "max-cut problem" defined as follows. Given a graph $G = (V, E)$, find a partition the vertex set V so that the number of edges crossing the partition is maximized. Here, a partition (V_1, V_2) of V consists of two sets V_1, V_2 satisfying $V_1 \cup V_2 = V$ and $V_1 \cap V_2 = \emptyset$, and the set of edges crossing the partition is basically $\{uv \in E : u \in V_1, v \in V_2\}$. For example, in Figure 19.1, there is a graph of 5 vertices partitioned into red and black vertices, and the edges highlighted are the ones crossing the partition. The max-cut problem is NP-hard, and we may consider approximation algorithms for it. There is a simple randomized algorithm that achieves a 1/2-approximation ratio as follows:

1. Initialize $V_1 = \emptyset$ and $V_2 = \emptyset$.

2. For each vertex $v \in V$, add it to V_1 or V_2 with probability 1/2 each.

Theorem 19.3. *Let $G = (V, E)$ be an undirected graph. The randomized algorithm above outputs a partition (V_1, V_2) such that the expected number of edges crossing the partition is $|E|/2$.*

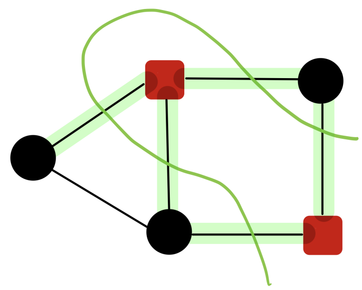

Figure 19.1: Edges crossing a partition

Proof. Note that

$$\mathbb{E}\left[|\delta(V_1, V_2)|\right] = \sum_{uv\in E} \mathbb{P}\left[uv \in \delta(V_1, V_2)\right].$$

Here, for each edge $uv \in E$, it holds that

$$\mathbb{P}\left[uv \in \delta(V_1, V_2)\right] = \mathbb{P}\left[(u \in V_1, v \in V_2) \text{ or } (u \in V_2, v \in V_1)\right] = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}.$$

Therefore, we have

$$\mathbb{E}\left[|\delta(V_1, V_2)|\right] \geq \sum_{uv\in E} \frac{1}{2} = \frac{1}{2}|E|,$$

as required. □

In fact, a simple deterministic greedy algorithm also achieves a 1/2-approximation ratio as follows:

1. Order and label the vertices as $1, \ldots, n$.
2. For each vertex j, consider $E_j = \{ij \in E : i < j\}$, i.e., the set of edges connecting vertex j and the previous vertices $1, \ldots, j-1$.
3. Initiallize $V_1 = \{1\}$ and $V_2 = \emptyset$.
4. For $j = 2, \ldots, n$, add vertex j to V_1 if the number of edges in E_j connecting to V_2 is larger than or equal to that connecting to V_1; otherwise, add it to V_2.

Theorem 19.4. *Let $G = (V, E)$ be an undirected graph. The greedy algorithm above outputs a partition (V_1, V_2) such that the number of edges crossing the partition is at least $|E|/2$.*

Proof. Note that $E_1, \ldots, E_n$ partition the edge set E. Hence, we have

$$|\delta(V_1, V_2)| = \sum_{j=1}^{n} |E_j \cap \delta(V_1, V_2)|.$$

For each j, by the greedy choice of adding vertex j to either V_1 or V_2, we have

$$|E_j \cap \delta(V_1, V_2)| \geq \frac{1}{2}|E_j|.$$

Combining these two inequalities yields

$$|\delta(V_1, V_2)| \geq \sum_{j=1}^{n} \frac{1}{2}|E_j| = \frac{1}{2}|E|,$$

as required. □

References

[Bac18] Francis Bach. Efficient algorithms for non-convex isotonic regression through submodular optimization. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. 2

[BMBK17] Andrew An Bian, Baharan Mirzasoleiman, Joachim Buhmann, and Andreas Krause. Guaranteed Non-convex Optimization: Submodular Maximization over Domains. In Aarti Singh and Jerry Zhu, editors, *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 54 of *Proceedings of Machine Learning Research*, pages 111–120, 20–22 Apr 2017. 19.1, 2, 19.2

[DJ12] Nikhil R. Devanur and Kamal Jain. Online matching with concave returns. In *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing*, STOC '12, page 137–144, 2012. 2

[EF16] Reza Eghbali and Maryam Fazel. Designing smoothing functions for improved worst-case competitive ratio in online optimization. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. 2

[MSVV05] A. Mehta, A. Saberi, U. Vazirani, and V. Vazirani. Adwords and generalized on-line matching. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 264–273, 2005. 2

[SJ17] Matthew Staib and Stefanie Jegelka. Robust budget allocation via continuous submodular functions. In Doina Precup and Yee Whye Teh, editors, *International Conference on Machine Learning (ICML)*, volume 70 of *Proceedings of Machine Learning Research*, pages 3230–3240, 06–11 Aug 2017. 2

[SKIK14] Tasuku Soma, Naonori Kakimura, Kazuhiro Inaba, and Ken-ichi Kawarabayashi. Optimal budget allocation: Theoretical guarantee and efficient algorithm. In *International Conference on Machine Learning (ICML)*, pages 351–359, 2014. 2