# Curvature and Optimal Algorithms for Learning and Minimizing Submodular Functions

Rishabh Iyer[1], Stefanie Jegelka[2], and Jeff Bilmes[1]

[1]University of Washington, Seattle; [2]University of California, Berkeley

## Overview

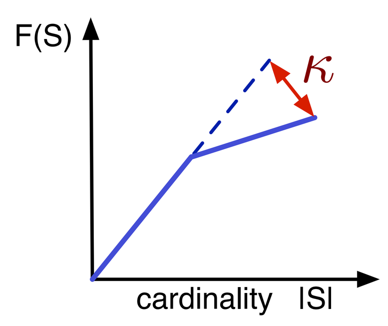
- Introduce the notion of *curvature*, to provide better connections between theory and practice.
- Study the role of curvature in:
  - Approximating submodular functions everywhere
  - Learning Submodular functions
  - Constrained Minimization of submodular functions.
- Provide improved curvature-dependent worst case approximation guarantees and matching hardness results

## Curvature of a Submodular function

- Define three variants of curvature of a monotone submodular function as:

$$\kappa_f = 1 - \min_{j \in V} \frac{f(j \mid V \setminus j)}{f(j)}, \quad \kappa_f(S) = 1 - \min_{j \in S} \frac{f(j|S\setminus j)}{f(j)}, \quad \hat{\kappa}_f(S) = 1 - \frac{\sum_{j \in S} f(j|S\setminus j)}{\sum_{j \in S} f(j)}$$

  - Proposition: $\hat{\kappa}_f(S) \leq \kappa_f(S) \leq \kappa_f$.
  - Captures the linearity of a submodular function.
  - A more gradual characterization of the hardness of various problems.
  - Investigated for submodular maximization (Conforti & Cornuejols, 1984).




## Main Ideas

- *Curve-Normalized form:* Given a monotone submodular function, the curve-normalized version of $f$ is:

$$f^\kappa(X) = \frac{f(X) - (1-\kappa_f)\sum_{j \in X} f(j)}{\kappa_f} \quad (1)$$

- Idea: Decompose $f$ as $f(X) = f_{\text{difficult}}(X) + m_{\text{easy}}(X)$ where $f_{\text{difficult}}(X) = \kappa_f f^\kappa(X)$ and $m_{\text{easy}}(X) = (1-\kappa_f)\sum_{j \in X} f(j)$.
- Lemma: If $f$ is monotone submodular, then $f^\kappa(X)$ is also monotone non-negative submodular function. Furthermore, $f^\kappa(X) \leq \sum_{j \in X} f(j)$.
- *Lower bounds:* Also show curvature-dependent lower bounds.

## Approximating Submodular functions Everywhere

**Problem:** Given a submodular function $f$ in form of a value oracle, find an approximation $\hat{f}$ (within polynomial time and space), such that $\hat{f}(X) \leq f(X) \leq \alpha_1(n)\hat{f}(X), \forall X \subseteq V$ for a polynomial $\alpha_1(n)$.

- We provide a blackbox technique to transform bounds into curvature dependent ones.
- Main technique: Approximate the curve-normalized version $f^\kappa$ as $\hat{f}^\kappa$, such that $\hat{f}^\kappa(X) \leq f^\kappa(X) \leq \alpha(n)\hat{f}^\kappa(X)$.

**Theorem**: The function $\hat{f}(X) \triangleq \kappa_f \hat{f}^\kappa(X) + (1-\kappa_f)\sum_{j \in X} f(j)$ satisfies

$$\hat{f}(X) \leq f(X) \leq \frac{\alpha(n)}{1 + (\alpha(n)-1)(1-\kappa_f)}\hat{f}(X) \leq \frac{\hat{f}(X)}{1-\kappa_f}. \quad (2)$$

- **Ellipsoidal Approximation:**
  - The Ellipsoidal Approximation algorithm of Goemans et al, provides a function of the form $\sqrt{w^f(X)}$ with an approximation factor of $\alpha_1(n) = O(\sqrt{n}\log n)$.
  - Corollary: There exists a function of the form, $f^{ea}(X) = \kappa_f\sqrt{w^{f^\kappa}(X)} + (1-\kappa_f)\sum_{j \in X} f(j)$ such that,

$$f^{ea}(X) \leq f(X) \leq O\left(\frac{\sqrt{n}\log n}{1 + (\sqrt{n}\log n - 1)(1-\kappa_f)}\right) f^{ea}(X). \quad (3)$$

  - Lower bound: Given a submodular function $f$ with curvature $\kappa_f$, there does not exist any polynomial-time algorithm that approximates $f$ within a factor of $\frac{n^{1/2-\epsilon}}{1+(n^{1/2-\epsilon}-1)(1-\kappa_f)}$, for any $\epsilon > 0$.
- **Modular Upper Bound:**
  - A simplest approximation (and upper bound) is $\hat{f}^m(X) = \sum_{j \in X} f(j)$.
  - Lemma: Given a monotone submodular function $f$, it holds that,

$$f(X) \leq \hat{f}^m(X) = \sum_{j \in X} f(j) \leq \frac{|X|}{1 + (|X|-1)(1-\hat{\kappa}_f(X))} f(X) \quad (4)$$

  - This bound is tight for the class of modular approximations.
  - Corollary: The class of functions, $f(X) = \sum_{i=1}^k \lambda_i [w_i(X)]^a, \lambda_i \geq 0$, satisfies $f(X) \leq \sum_{j \in X} f(j) \leq |X|^{1-a} f(X)$.

## Learning Submodular Functions

**Problem:** Given i.i.d training samples $\{(X_i, f(X_i))\}_{i=1}^m$ from a distribution $\mathcal{D}$, learn an approximation $\hat{f}(X)$ that is, with probability $1-\delta$, within a multiplicative factor of $\alpha_2(n)$ from $f$.

- Balcan & Harvey propose an algorithm which PMAC learns any submodular function upto a factor of $\sqrt{n+1}$.
- We improve this bound to a curvature dependent one.

**Lemma:** Let $f$ be a monotone submodular function for which we know an upper bound on its curvature $\kappa_f$ and the singleton weights $f(j)$ for all $j \in V$. There is an poly-time algorithm which PMAC-learns $f$ within a factor of $\frac{\sqrt{n+1}}{1+(\sqrt{n+1}-1)(1-\kappa_f)}$.

- We also provide an algorithm which does not need the singleton weights.

**Lemma:** If $f$ is a monotone submodular function with known curvature (or a known upper bound) $\hat{\kappa}_f(X), \forall X \subseteq V$, then for every $\epsilon, \delta > 0$ there is an algorithm which PMAC learns $f(X)$ within a factor of $1 + \frac{|X|}{1+(|X|-1)(1-\hat{\kappa}_f(X))}$.

- Corollary: The class of functions $f(X) = \sum_{i=1}^k \lambda_i [w_i(X)]^a, \lambda_i \geq 0$, can be learnt to a factor of $|X|^{1-a}$.
- Lower bound: Given a class of submodular functions with curvature $\kappa_f$, there does not exist a polynomial-time algorithm that is guaranteed to PMAC-learn $f$ within a factor of $\frac{n^{1/3-\epsilon'}}{1+(n^{1/3-\epsilon'}-1)(1-\kappa_f)}$, for any $\epsilon' > 0$.

## Constrained Submodular Minimization

**Problem:** Minimize a submodular function $f$ over a family $\mathcal{C}$ of feasible sets, i.e., $\min_{X \in \mathcal{C}} f(X)$. $\mathcal{C}$ could be constraints of the form cardinality (knapsack) constraints, cuts, paths, matchings, trees etc.

- Main framework is to choose a surrogate function $\hat{f}$, and optimize it instead of $f$.
- **Ellipsoidal Approximation based (EA):**
  - Use the curvature based Ellipsoidal Approximation as the surrogate function.
  - Lemma: For a submodular function with curvature $\kappa_f < 1$, algorithm EA will return a solution $\widehat{X}$ that satisfies

$$f(\widehat{X}) \leq O\left(\frac{\sqrt{n}\log n}{(\sqrt{n}\log n - 1)(1-\kappa_f) + 1}\right) f(X^*).$$

- **Modular Upper bound based:**
  - Use the simple modular upper bound as a surrogate.
  - Lemma: Let $\hat{X} \in \mathcal{C}$ be the solution for minimizing $\sum_{j \in X} f(j)$ over $\mathcal{C}$. Then

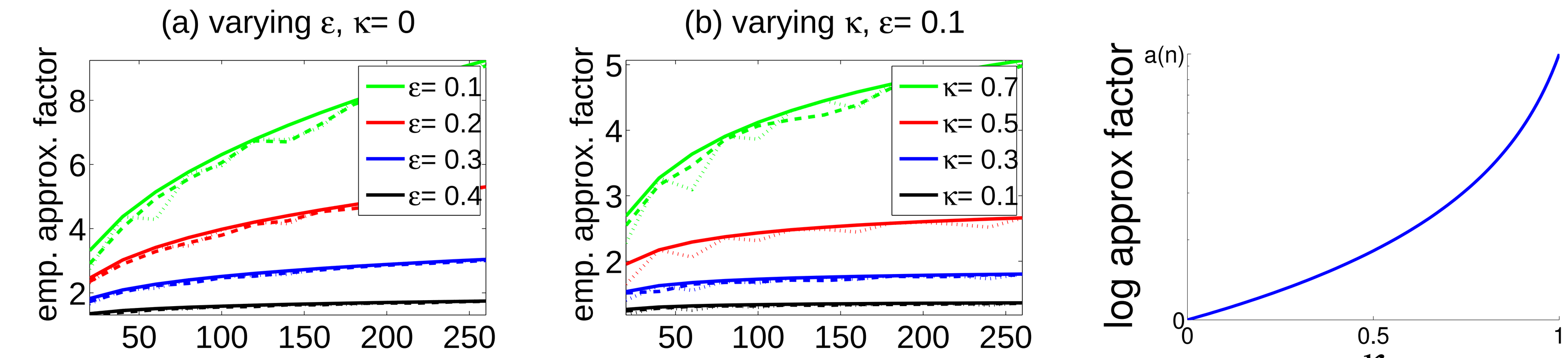
$$f(\hat{X}) \leq \frac{|X^*|}{1 + (|X^*|-1)(1-\kappa_f(X^*))} f(X^*). \quad (5)$$

  - Corollary: The class of functions, $f(X) = \sum_{i=1}^k \lambda_i [w_i(X)]^a, \lambda_i \geq 0$, can be minimized upto a factor of $|X^*|^{1-a}$.

| Constraint | MUB | EA | Curvature-Ind. | Lower bound |
|---|---|---|---|---|
| Card. LB | $\frac{k}{1+(k-1)(1-\kappa_f)}$ | $O(\frac{\sqrt{n}\log n}{1+(\sqrt{n}\log n-1)(1-\kappa_f)})$ | $\theta(n^{1/2})$ | $\tilde{\Omega}(\frac{\sqrt{n}}{1+(\sqrt{n}-1)(1-\kappa_f)})$ |
| Spanning Tree | $\frac{n}{1+(n-1)(1-\kappa_f)}$ | $O(\frac{\sqrt{m}\log m}{1+(\sqrt{m}\log m-1)(1-\kappa_f)})$ | $\theta(n)$ | $\tilde{\Omega}(\frac{n}{1+(n-1)(1-\kappa_f)})$ |
| Matchings | $\frac{n}{2+(n-2)(1-\kappa_f)}$ | $O(\frac{\sqrt{m}\log m}{1+(\sqrt{m}\log m-1)(1-\kappa_f)})$ | $\theta(n)$ | $\tilde{\Omega}(\frac{n}{1+(n-1)(1-\kappa_f)})$ |
| s-t path | $\frac{n}{1+(n-1)(1-\kappa_f)}$ | $O(\frac{\sqrt{m}\log m}{1+(\sqrt{m}\log m-1)(1-\kappa_f)})$ | $\theta(n^{2/3})$ | $\tilde{\Omega}(\frac{n^{2/3}}{1+(n^{2/3}-1)(1-\kappa_f)})$ |
| s-t cut | $\frac{m}{1+(m-1)(1-\kappa_f)}$ | $O(\frac{\sqrt{m}\log m}{1+(\log m\sqrt{m}-1)(1-\kappa_f)})$ | $\theta(\sqrt{n})$ | $\tilde{\Omega}(\frac{\sqrt{n}}{1+(\sqrt{n}-1)(1-\kappa_f)})$ |

Table : Summary of our results for constrained minimization.

- **Effect of Curvature:** Polynomial change in the bounds!
- **Experiments:**
  - Define a function $f_R(X) = \kappa \min\{|X \cap \bar{R}| + \beta, |X|, \alpha\} + (1-\kappa)|X|$.
  - Choose $\alpha = n^{1/2+\epsilon}$ and $\beta = n^{2\epsilon}$, and $\mathcal{C} = \{X : |X| \geq \alpha\}$.




## Acknowledgements

Based upon work supported by National Science Foundation Grant No. IIS-1162606, and by a Google, a Microsoft, and an Intel research award. This was also funded in part by Office of Naval Research under grant no. N00014-11-1-0688, NSF CISE Expeditions award CCF-1139158, DARPA XData Award FA8750-12-2-0331, and gifts from Amazon Web Services, Google, SAP, Blue Goji, Cisco, Clearstory Data, Cloudera, Ericsson, Facebook, General Electric, Hortonworks, Intel, Microsoft, NetApp, Oracle, Samsung, Splunk, VMware and Yahoo!