Fast Decomposable Submodular Function Minimization using Constrained Total Variation

K S Sesh Kumar [1]
Francis Bach [2] Thomas Pock [3]

[1]Data Science Institute
Imperial College London, UK
s.karri@imperial.ac.uk

[2]INRIA and Ecole normale superieure
PSL Research University, Paris France.
francis.bach@inria.fr

[3]Institute of Computer Graphics and Vision,
Graz University of Technology, Graz, Austria.
pock@icg.tugraz.at

Neural Information Processing Systems, 2019.

Submodular functions

- **Definition**: $F : 2^V \to \mathbb{R}$ is **submodular** if and only if

$$\forall A, B \subseteq V, F(A) + F(B) \geq F(A \cup B) + F(A \cap B)$$

  - Equality for *modular* functions
  - Always assume $F(\varnothing) = 0$.

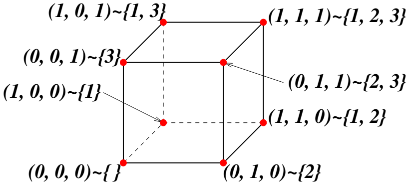
- **Equivalent definition** :

$$\forall k \in V, A \to F(A \cup \{k\}) - F(A) \text{ is non increasing}$$

  - Diminishing returns property

Subsets as pseudo boolean functions

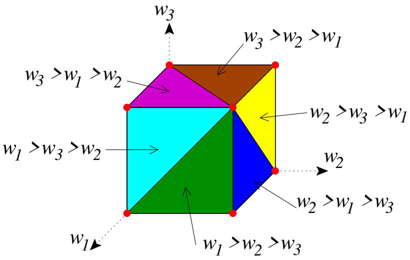
- Subsets may be identified with elements of $\{0, 1\}^p$.



Lovász extension

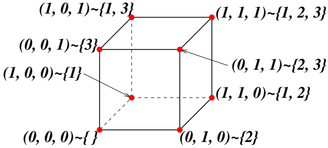
- Given any set-function F and w such that $w_{j_1} \geqslant \cdots \geqslant w_{j_p}$,

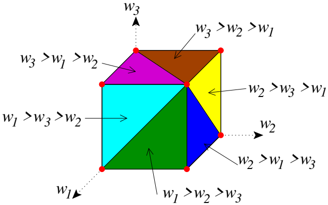
$$f(w) = \sum_{k=1}^{p-1} (w_{j_k} - w_{j_{k+1}}) F(\{j_1, \ldots, j_k\}) + w_{j_p} F(\{j_1, \ldots, j_p\})$$



  - f is piecewise-linear and positively homogeneous
  - if $w = 1_A, f(w) = F(A)$
  - Extension from $\{0,1\}^p$ to $\mathbb{R}^p$
- **Theorem**: F is submodular if and only if f is convex

Cut functions and Lovász extension

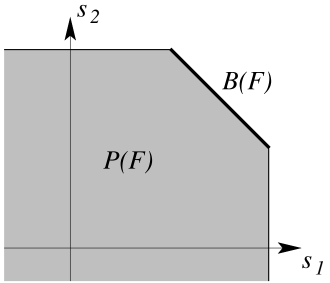


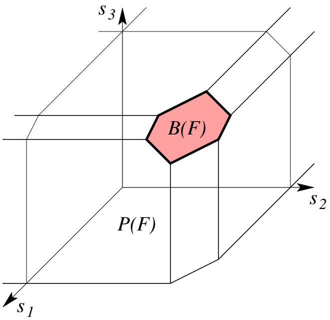


- **Cut function.** $F : 2^V \to \mathbb{R}$ such that $F(A) = \sum_{i \in V} \sum_{j \in \mathcal{N}(i)} a_{ij} |1_{i \in A} - 1_{j \in A}|$.
- **Lovász extension.** $f : [0,1]^p \to \mathbb{R}$ such that $f(w) = \sum_{i \in V} \sum_{j \in \mathcal{N}(i)} a_{ij} |w_i - w_j|$.
- If F is a <u>cut function</u> then f is the corresponding <u>total variation</u> problem.

Submodular Functions and base polyhedra

- Submodular polyhedron:
  $P(F) = \{ s \in \mathbb{R}^p, \ \forall A \subset V, \ s(A) \leqslant F(A) \}$
- Base polyhedron: $B(F) = P(F) \cap \{ s(V) = F(V) \}$
- Many facets (up to 2^p), many extreme points (up to $p!$)





Lovász extension and base polyhedra

- **Fundamental property**: If F is submodular, maximizing linear functions may be done by a "greedy algorithm"
  - Let $w \in \mathbb{R}_+^p$ such that $w_{j_1} \geqslant \cdots \geqslant w_{j_p}$
  - Let $s_{j_k} = F(\{j_1, \dots, j_k\}) - F(\{j_1, \dots, j_{k-1}\})$ for $k \in \{1, \dots, p\}$
  - Then $f(w) = \max_{s \in P(F)} w^\top s = \max_{s \in B(F)} w^\top s$
- **Representation of $f(w)$ as a support function**:

$$\boxed{f(w) = \max_{s \in B(F)} w^\top s}$$

From SFM to continuous minimization problems

- **Related optimization problems**

  (SFM**D**) Discrete $\min_{A \subset V} F(A) - u(A)$[1]

  (SFM**C**) Continuous $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \sum_{j=1}^{n} \psi(w_j)$

- **Solving (SFMC) is equivalent to**

$$\min_{A \subset V} F(A) - u(A) + |A| \psi'(\alpha), \forall \alpha \in \mathbb{R}$$

  - $\arg\min_{A \subseteq V} F(A) - u(A) + |A|\psi'(\alpha) = \{k, w_k^* \geqslant \alpha\}$
  - Consequence of F being submodular.
  - $\alpha = 0$ solves (SFM**D**).

- If (SFM**D**) is graph cut $\Rightarrow$ (SFM**C**) is parametric maxflow.

[1] $u(A) = u^\top 1_A$

Overview

- **Goal**

  $(\mathrm{SFM}\mathbf{D})$ Discrete $\quad \min_{A \subset V} \sum_{i=1}^{r} F_i(A) - u(A)$

- **Assuming the oracles**

  $(\mathrm{SFM}\mathbf{D}_\mathrm{i})$ Discrete $\quad \min_{A \subset V} F_i(A) - t(A)$

- **Define**
  - $(\mathrm{SFM}\mathbf{C})$ $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \sum_{j=1}^{n} \psi(w_j)$
  - $(\mathrm{SFM}\mathbf{C}_\mathrm{i})$ $\min_{w \in \mathbb{R}^p} f_i(w) - t^\top w + \sum_{j=1}^{n} \psi(w_j)$
  - Define ψ

$$\psi(w) = \begin{cases} \frac{1}{2} w^2 & \text{if } |w| \leqslant \varepsilon, \\ +\infty & \text{otherwise}, \end{cases} \tag{1}$$

- **Approach**
  - Solve $(\mathrm{SFM}\mathbf{C}_\mathrm{i})$ using $(\mathrm{SFM}\mathbf{D}_\mathrm{i})$
  - Solve $(\mathrm{SFM}\mathbf{D})$ by solving $(\mathrm{SFM}\mathbf{C})$ using $(\mathrm{SFM}\mathbf{C}_\mathrm{i})$