Convex Optimization for Parallel Energy Minimization

K. S. Sesh Kumar

Data Science Institute
Imperial College London

October 1, 2019

Overview

Submodularity and Examples

Notations and Definitions

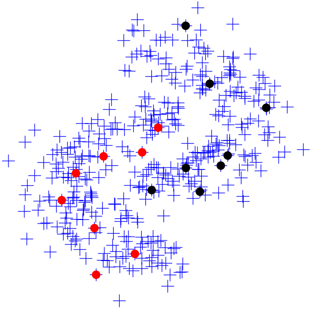
SFM and the corresponding smooth problems

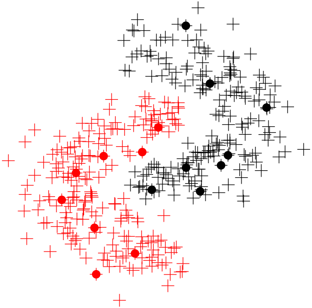
Active set methods for submodular minimization

Active set methods for parallel submodular minimization

Results

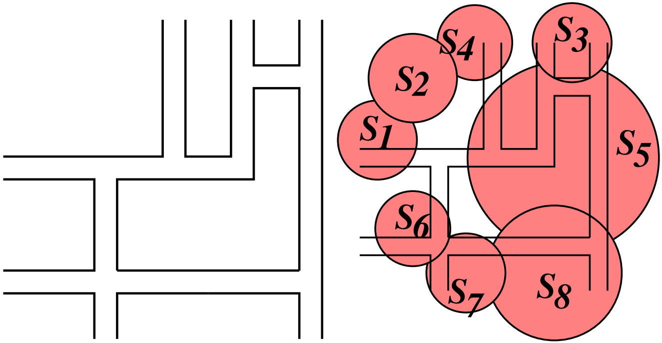
Semi - supervised clustering





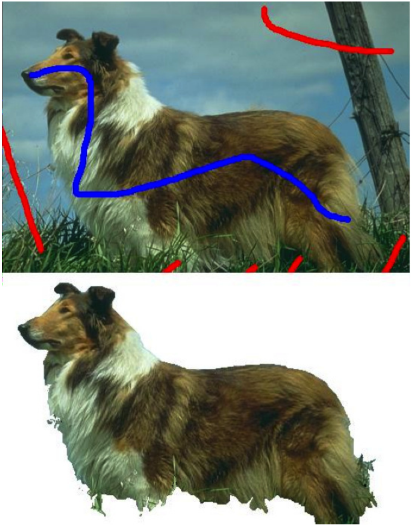
- Submodular function minimization

Sensor placement



- Krause and Guestrin, 2005.
- Submodular function maximization

Energy Minimization in Computer Vision



- Graph cuts and image segmentation

Total Variation Denoising

- Chambolle, 2005.

Submodular functions

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

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

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 $F(\varnothing) = 0$.

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** :

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

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 $F(\varnothing) = 0$.

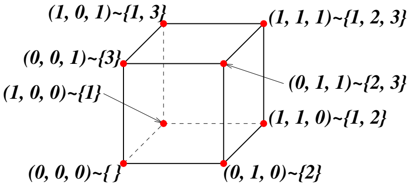
- **Equivalent definition** :

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

  - Diminishing returns property

Subsets as pseudo boolean functions

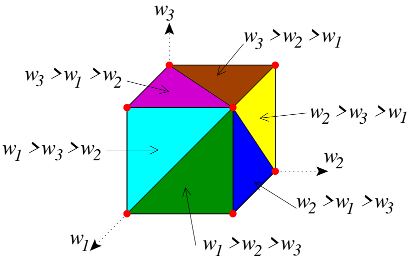
- Subsets may be identified with elements of $\{0,1\}^p$.



Lovász extension

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

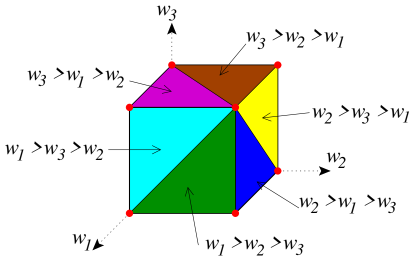
$$f(w) = \sum_{k=1}^{p-1} (w_{j_k} - w_{j_{k+1}}) F(\{j_1, \ldots, j_k\}) + w_{j_p} F(\{j_1, \ldots, j_p\})$$



Lovász extension

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

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

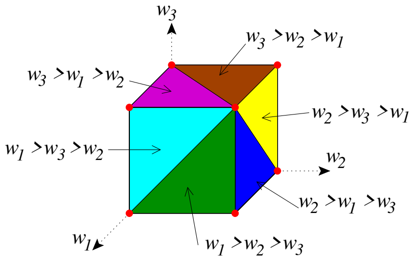


  - f is piecewise-linear and positively homogeneous
  - if $w = 1_A$, $f(w) = F(A)$
  - Extension from $\{0,1\}^p$ to $\mathbb{R}^p$

Lovász extension

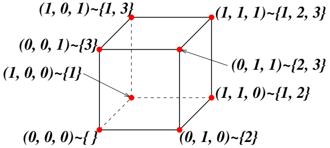
- Given any set-function F and w such that $w_{j_1} \geqslant \cdots \geqslant w_{j_p}$,

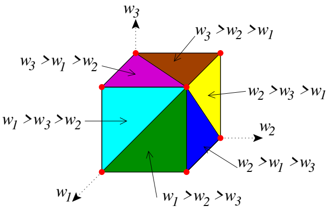
$$f(w) = \sum_{k=1}^{p-1} (w_{j_k} - w_{j_{k+1}}) F(\{j_1, \ldots, j_k\}) + w_{j_p} F(\{j_1, \ldots, j_p\})$$



  - 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

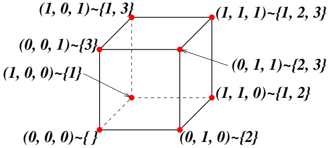
Cut functions and Lovász extension

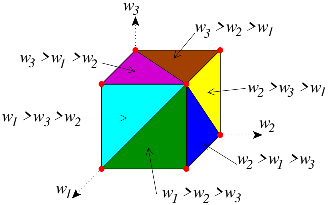




- **Cut function.** $F : 2^V \to \mathbb{R}$ such that $F(A) = \sum_{i \in V} \sum_{j \in \mathcal{N}(i)} a_{ij} |1_{i \in A} - 1_{j \in A}|$.

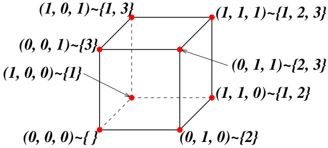
Cut functions and Lovász extension

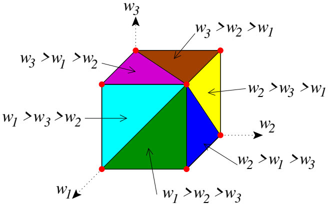




- **Cut function.** $F : 2^V \to \mathbb{R}$ such that $F(A) = \sum_{i \in V} \sum_{j \in \mathcal{N}(i)} a_{ij} |1_{i \in A} - 1_{j \in A}|$.
- **Lovász extension.** $f : [0,1]^p \to \mathbb{R}$ such that $f(w) = \sum_{i \in V} \sum_{j \in \mathcal{N}(i)} a_{ij} |w_i - w_j|$.

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 cut function then f is the corresponding total variation problem.

From submodular minimization to smooth problems

- **Related optimization problems**

(D) Discrete $\min_{A \subset V} F(A) - u(A)^1 = \min_{w \in \{0,1\}^p} f(w) - u^\top w$

(C) Continuous $\min_{w \in [0,1]^p} f(w) - u^\top w$

(S) Smooth $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$

---

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

From submodular minimization to smooth problems

- **Related optimization problems**

  (D) Discrete $\min_{A \subset V} F(A) - u(A)^1 = \min_{w \in \{0,1\}^p} f(w) - u^\top w$

  (C) Continuous $\min_{w \in [0,1]^p} f(w) - u^\top w$

  (S) Smooth $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$

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

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

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

---

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

From submodular minimization to smooth problems

- **Related optimization problems**

  (D) Discrete $\min_{A \subset V} F(A) - u(A)$[1] $= \min_{w \in \{0,1\}^p} f(w) - u^\top w$

  (C) Continuous $\min_{w \in [0,1]^p} f(w) - u^\top w$

  (S) Smooth $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$

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

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

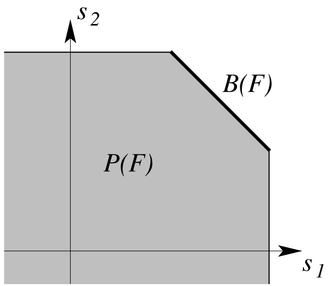
  - $\arg\min_{A \subseteq V} F(A) - u(A) + \lambda |A| = \{k, w_k \geqslant \lambda\}$
  - Consequence of F being submodular.
  - $\lambda = 0$ solves **(D)**.

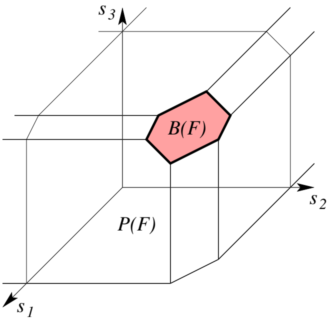
- If **(D)** is graph cut $\Rightarrow$ **(S)** is parametric maxflow.

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

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 $s_{j_k} = F(\{j_1, \ldots, j_k\}) - F(\{j_1, \ldots, j_{k-1}\})$ for $k \in \{1, \ldots, p\}$
  - Then $f(w) = \max_{s \in P(F)} w^\top s = \max_{s \in B(F)} w^\top s$

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 $s_{j_k} = F(\{j_1, \ldots, j_k\}) - F(\{j_1, \ldots, j_{k-1}\})$ for $k \in \{1, \ldots, p\}$
  - 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**:

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

Lovász extension and base polyhedra

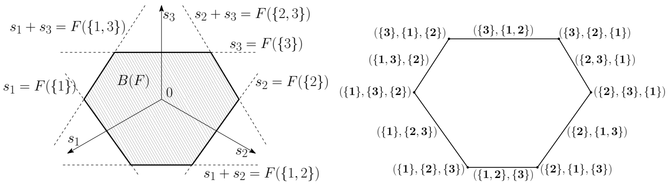
- **Fundamental property**: If F is submodular, maximizing linear functions may be done by a "greedy algorithm"
  - Let $w \in \mathbb{R}_+^p$ such that $w_{j_1} \geqslant \cdots \geqslant w_{j_p}$
  - Let $s_{j_k} = F(\{j_1, \ldots, j_k\}) - F(\{j_1, \ldots, j_{k-1}\})$ for $k \in \{1, \ldots, p\}$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

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

- **Primal - Dual**
  - **Primal** : $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$
  - **Dual** : $\max_{s \in B(F)} -\frac{1}{2}\|u - s\|_2^2$
  - At optimal, $w^* = u - s^*$

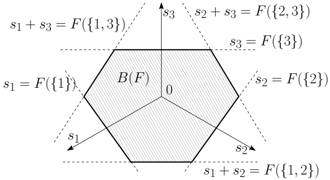
Faces of the Base Polyhedon and Ordered Partitions

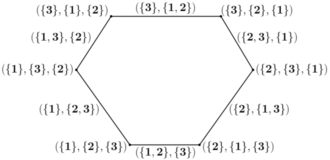
$$B(F) = \left\{ s \in \mathbb{R}^p,\ \forall A \subset V,\ s(A) \leqslant F(A), s(V) = F(V) \right\}$$



Faces of the Base Polyhedon and Ordered Partitions

$$B(F) = \left\{ s \in \mathbb{R}^p,\ \forall A \subset V,\ s(A) \leqslant F(A), s(V) = F(V) \right\}$$

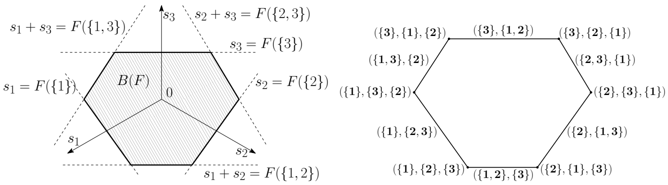




- Given an Ordered Partition $\mathcal{A} = (A_1, \ldots, A_m)$
- Let $B_i = A_1 \cup \ldots \cup A_i, \forall i = 1, \ldots, m$. $B_m = V$

Faces of the Base Polyhedon and Ordered Partitions

$$B(F) = \left\{ s \in \mathbb{R}^p,\ \forall A \subset V,\ s(A) \leqslant F(A), s(V) = F(V) \right\}$$

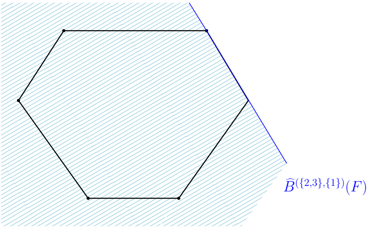


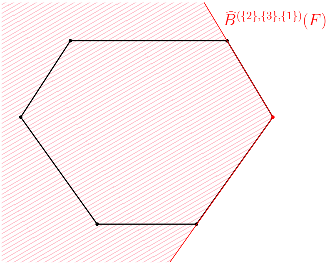
- Given an Ordered Partition $\mathcal{A} = (A_1, \ldots, A_m)$
- Let $B_i = A_1 \cup \ldots \cup A_i, \forall i = 1, \ldots, m$. $B_m = V$
- Outer approximation of $B(F)$ by the ordered partition $\mathcal{A}$

$$\widehat{B}^{\mathcal{A}}(F) = \{s \in \mathbb{R}^n, s(V) = F(V), \forall i = \{1, \ldots, m{-}1\}, s(B_i) \leq F(B_i)\}$$

Outer Approximation of Base Polytope and Optimization

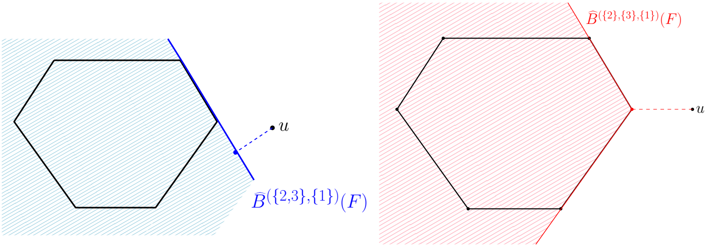
$$\widehat{B}^{\mathcal{A}}(F) = \{s \in \mathbb{R}^n, s(V) = F(V), \forall i = \{1, \ldots, m{-}1\}, s(B_i) \leq F(B_i)\}$$





Outer Approximation of Base Polytope and Optimization

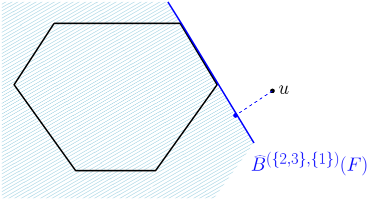
$$\widehat{B}^{\mathcal{A}}(F) = \{s \in \mathbb{R}^n, s(V) = F(V), \forall i = \{1, \ldots, m{-}1\}, s(B_i) \leq F(B_i)\}$$

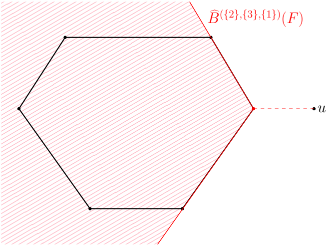


- $\max_{s \in \widehat{B}^{\mathcal{A}}(F)} -\frac{1}{2}\|u - s\|_2^2$

Outer Approximation of Base Polytope and Optimization

$$\widehat{B}^{\mathcal{A}}(F) = \{s \in \mathbb{R}^n, s(V) = F(V), \forall i = \{1, \dots, m{-}1\}, s(B_i) \leq F(B_i)\}$$

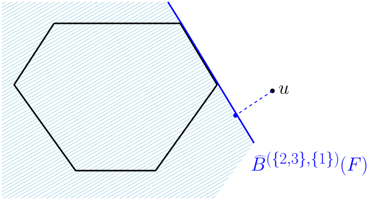


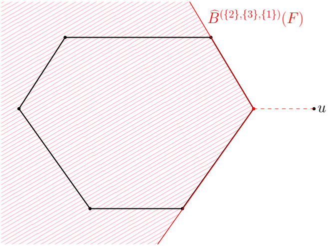


- $\max_{s \in \widehat{B}^{\mathcal{A}}(F)} -\frac{1}{2}\|u - s\|_2^2$
- $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$, s.t. w is compatible with $\mathcal{A}$

Outer Approximation of Base Polytope and Optimization

$$\widehat{B}^{\mathcal{A}}(F) = \{s \in \mathbb{R}^n, s(V) = F(V), \forall i = \{1, \dots, m-1\}, s(B_i) \leq F(B_i)\}$$

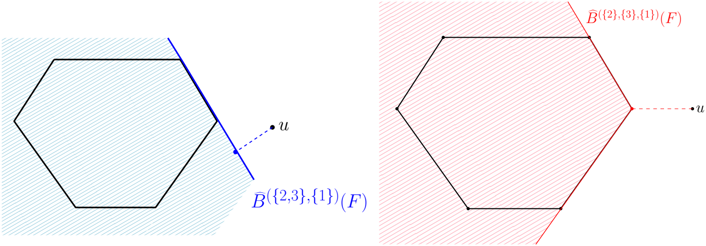




- $\max_{s \in \widehat{B}^{\mathcal{A}}(F)} -\frac{1}{2}\|u - s\|_2^2$
- $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$, s.t. w is compatible with $\mathcal{A}$
  - $w = \sum_{i=1}^m v_i 1_{A_i}$
  - $f(w) = \sum_{i=1}^m v_i [F(B_i) - F(B_{i-1})]$

Outer Approximation of Base Polytope and Optimization

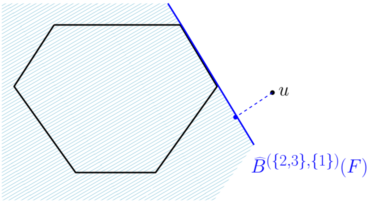
$$\widehat{B}^{\mathcal{A}}(F) = \{s \in \mathbb{R}^n, s(V) = F(V), \forall i = \{1, \ldots, m-1\}, s(B_i) \le F(B_i)\}$$

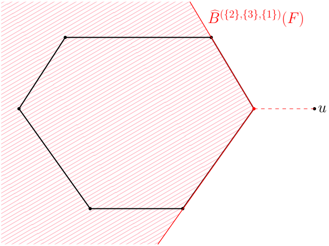


- $\max_{s \in \widehat{B}^{\mathcal{A}}(F)} -\frac{1}{2}\|u - s\|_2^2$
- $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$, s.t. w is compatible with $\mathcal{A}$
  - $w = \sum_{i=1}^m v_i 1_{A_i}$
  - $f(w) = \sum_{i=1}^m v_i [F(B_i) - F(B_{i-1})]$
- $\min_{v \in \mathbb{R}^m} \sum_{i=1}^m v_i [F(B_i) - F(B_i - 1) - u(A_i)] + \frac{1}{2} \sum_{i=1}^m |A_i| v_i^2$ such that $v_1 \ge \ldots \ge v_m$

Outer Approximation of Base Polytope and Optimization

$$\widehat{B}^{\mathcal{A}}(F) = \{s \in \mathbb{R}^n, s(V) = F(V), \forall i = \{1, \ldots, m-1\}, s(B_i) \leq F(B_i)\}$$





- $\max_{s \in \widehat{B}^{\mathcal{A}}(F)} -\frac{1}{2}\|u - s\|_2^2$
- $\min_{w \in \mathbb{R}^p} f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$, s.t. w is compatible with $\mathcal{A}$
  - $w = \sum_{i=1}^m v_i 1_{A_i}$
  - $f(w) = \sum_{i=1}^m v_i[F(B_i) - F(B_{i-1})]$
- $\min_{v \in \mathbb{R}^m} \sum_{i=1}^m v_i[F(B_i) - F(B_i - 1) - u(A_i)] + \frac{1}{2}\sum_{i=1}^m |A_i| v_i^2$ such that $v_1 \geq \ldots \geq v_m$
- Isotonic regression solved using wpav, $s = u - w$
- Refines $\mathcal{A}$ by merging partitions and ensuring $v_1 > \ldots > v_{m'}$

Optimality test and Splitting partitions

- $s \in B(F)$, i.e., $\forall A \subset V, s(A) \leq F(A)$

Optimality test and Splitting partitions

- $s \in B(F)$, i.e., $\forall A \subset V, s(A) \leq F(A)$
- SFM oracle which solves

$$C = \min_{A \subset V} F(A) - s(A)$$

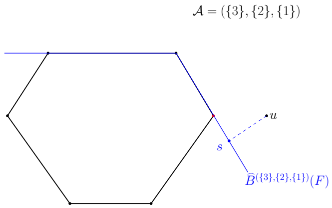
Optimality test and Splitting partitions

- $s \in B(F)$, i.e., $\forall A \subset V, s(A) \leq F(A)$
- SFM oracle which solves

$$C = \min_{A \subset V} F(A) - s(A)$$

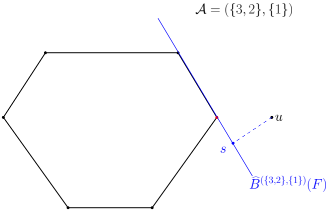
- If $F(C) - s(C) < 0$ then split all the ordered partitions of $\mathcal{A}$ that intersect with C

Active Set Algorithm



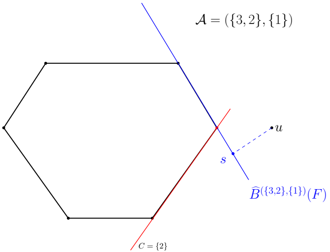
- Input : Submodular function F with SFM oracle, $u \in \mathbb{R}^p$, ordered partition $\mathcal{A}$
- Algorithm : iterate until convergence
  (a) Project u onto outer-approximation $\widehat{B}^{\mathcal{A}}(F)$ (using isotonic regression)

Active Set Algorithm



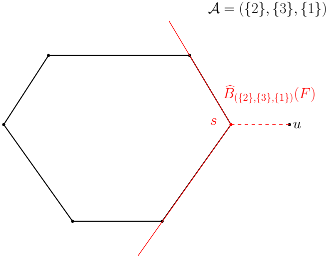
- Input : Submodular function F with SFM oracle, $u \in \mathbb{R}^p$, ordered partition $\mathcal{A}$
- Algorithm : iterate until convergence
  (a) Project u onto outer-approximation $\widehat{B}^{\mathcal{A}}(F)$ (using isotonic regression)
  (b) Refine partitions by merging partitions with equal values

Active Set Algorithm



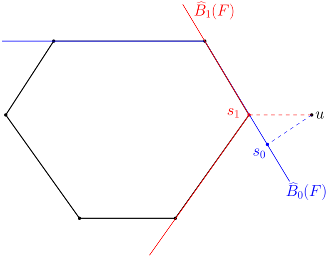
- Input : Submodular function F with SFM oracle, $u \in \mathbb{R}^p$, ordered partition $\mathcal{A}$
- Algorithm : iterate until convergence
  (a) Project u onto outer-approximation $\widehat{B}^{\mathcal{A}}(F)$ (using isotonic regression)
  (b) Refine partitions by merging partitions with equal values
  (c) Check for optimality of s by SFM oracle to solve $C = \min_{A \in V} F(C) - s(C)$

Active Set Algorithm



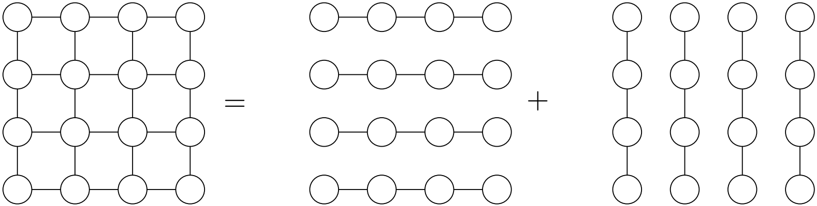
- Input : Submodular function F with SFM oracle, $u \in \mathbb{R}^p$, ordered partition $\mathcal{A}$
- Algorithm : iterate until convergence
  (a) Project u onto outer-approximation $\widehat{B}^{\mathcal{A}}(F)$ (using isotonic regression)
  (b) Refine partitions by merging partitions with equal values
  (c) Check for optimality of s by SFM oracle to solve $C = \min_{A \in V} F(C) - s(C)$
  (d) If not optimal then split the ordered partitions of $\mathcal{A}$ that intersect with C and goto (a).

Active Set Algorithm



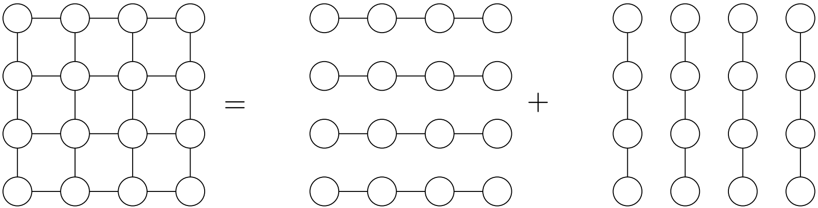
- Input : Submodular function F with SFM oracle, $u \in \mathbb{R}^p$, ordered partition $\mathcal{A}$
- Algorithm : iterate until convergence
  (a) Project u onto outer-approximation $\widehat{B}^{\mathcal{A}}(F)$ (using isotonic regression)
  (b) Refine partitions by merging partitions with equal values
  (c) Check for optimality of s by SFM oracle to solve $C = \min_{A \in V} F(C) - s(C)$
  (d) If not optimal then split the ordered partitions of $\mathcal{A}$ that intersect with C and goto (a).

Decomposable functions



- **Goal:** Use simpler SFM oracles to solve (S) and as a consequence solve (D)

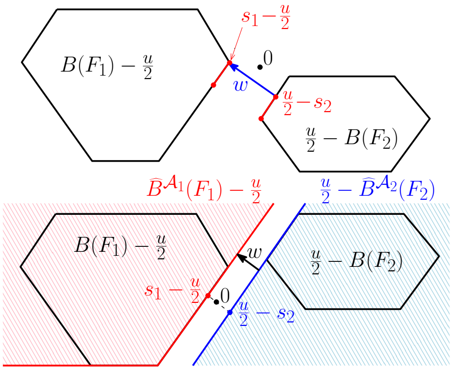
Decomposable functions



- **Goal:** Use simpler SFM oracles to solve (S) and as a consequence solve (D)
- **(D)** : $\min_{A \subset V} F_1(A) + F_2(A) - u(A)$
- **(S)** : $\min_{w \in \mathbb{R}^p} f_1(w) + f_2(w) - u^\top w + \frac{1}{2}\|w\|_2^2$
- **Dual** : $\min_{s_1 \in B(F_1) - u/2, -s_2 \in u/2 - B(F_2)} \|s_1 - (-s_2)\|_2$

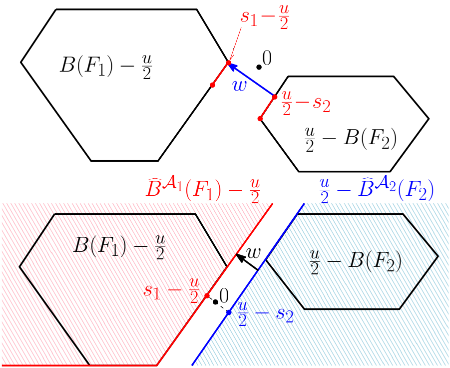
Closest points between two polytopes

Dual : $\min_{s_1 \in B(F_1)-u/2,\, s_2 \in u/2-B(F_2)} \|s_1 - s_2\|_2$



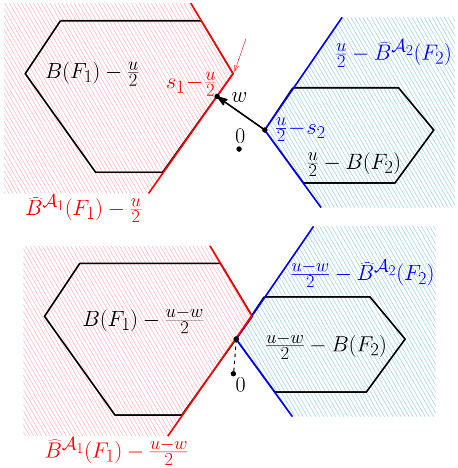
Closest points between two polytopes

Dual : $\min_{s_1 \in B(F_1)-u/2, s_2 \in u/2-B(F_2)} \|s_1 - s_2\|_2$



- Reflection methods for user-friendly submodular optimization. S.Jegelka, F. Bach and S. Sra, NIPS-2013.
- Uses expensive TV oracles.

Translated intersecting polytopes



Active set method for decomposable functions

- **Input**: Submodular function F_1 and F_2 with SFM oracles, $u \in \mathbb{R}^n$, ordered partitions $\mathcal{A}_1, \mathcal{A}_2$
- **Algorithm**: iterate until convergence (i.e., $\varepsilon_1 + \varepsilon_2$ small enough)

Active set method for decomposable functions

- **Input**: Submodular function F_1 and F_2 with SFM oracles, $u \in \mathbb{R}^n$, ordered partitions $\mathcal{A}_1, \mathcal{A}_2$
- **Algorithm**: iterate until convergence (i.e., $\varepsilon_1 + \varepsilon_2$ small enough)
  (a) Find $\mathcal{A} = \mathrm{coalesce}(\mathcal{A}_1, \mathcal{A}_2)$ and run isotonic regression to minimize $f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$ such that w is compatible with $\mathcal{A}$.

Active set method for decomposable functions

- **Input**: Submodular function F_1 and F_2 with SFM oracles, $u \in \mathbb{R}^n$, ordered partitions $\mathcal{A}_1, \mathcal{A}_2$
- **Algorithm**: iterate until convergence (i.e., $\varepsilon_1 + \varepsilon_2$ small enough)
  (a) Find $\mathcal{A} = \mathrm{coalesce}(\mathcal{A}_1, \mathcal{A}_2)$ and run isotonic regression to minimize $f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$ such that w is compatible with $\mathcal{A}$.
  (b) Run accelerated Dykstra's algorithm to find the projection of 0 onto the intersection of $\widehat{B}^{\mathcal{A}_1}(F_1) - u/2 + w/2$ and $u/2 - w/2 - \widehat{B}^{\mathcal{A}_2}(F_2)$.

Active set method for decomposable functions

- **Input**: Submodular function F_1 and F_2 with SFM oracles, $u \in \mathbb{R}^n$, ordered partitions $\mathcal{A}_1, \mathcal{A}_2$
- **Algorithm**: iterate until convergence (i.e., $\varepsilon_1 + \varepsilon_2$ small enough)
  (a) Find $\mathcal{A} = \mathrm{coalesce}(\mathcal{A}_1, \mathcal{A}_2)$ and run isotonic regression to minimize $f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$ such that w is compatible with $\mathcal{A}$.
  (b) Run accelerated Dykstra's algorithm to find the projection of 0 onto the intersection of $\widehat{B}^{\mathcal{A}_1}(F_1) - u/2 + w/2$ and $u/2 - w/2 - \widehat{B}^{\mathcal{A}_2}(F_2)$.
  (c) Merge the sets in $\mathcal{A}_j$ which are tight for s_j, $j \in \{1, 2\}$.

Active set method for decomposable functions

- **Input**: Submodular function F_1 and F_2 with SFM oracles, $u \in \mathbb{R}^n$, ordered partitions $\mathcal{A}_1, \mathcal{A}_2$
- **Algorithm**: iterate until convergence (i.e., $\varepsilon_1 + \varepsilon_2$ small enough)
  (a) Find $\mathcal{A} = \mathrm{coalesce}(\mathcal{A}_1, \mathcal{A}_2)$ and run isotonic regression to minimize $f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$ such that w is compatible with $\mathcal{A}$.
  (b) Run accelerated Dykstra's algorithm to find the projection of 0 onto the intersection of $\widehat{B}^{\mathcal{A}_1}(F_1) - u/2 + w/2$ and $u/2 - w/2 - \widehat{B}^{\mathcal{A}_2}(F_2)$.
  (c) Merge the sets in $\mathcal{A}_j$ which are tight for s_j, $j \in \{1, 2\}$.
  (d) Check optimality by solving $\min_{C_{j,i_j} \subseteq A_{j,i_j}} F_j(B_{j,i_j-1} \cup C_{j,i_j}) - F_j(B_{j,i+j-1}) - s_j(C_{j,i_j})$ for $i_j \in \{1, \ldots, m_j\}$, Monitor ε_1 and ε_2 such that $F_j(C_j) - s_j(C_j) \geqslant -\varepsilon_j$, $j = 1, 2$.

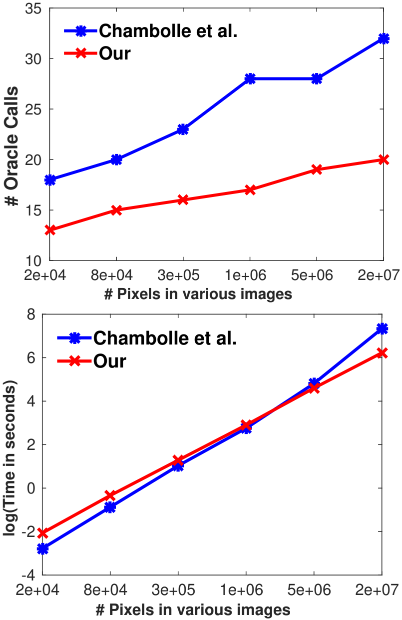
Active set method for decomposable functions

- **Input**: Submodular function F_1 and F_2 with SFM oracles, $u \in \mathbb{R}^n$, ordered partitions $\mathcal{A}_1, \mathcal{A}_2$
- **Algorithm**: iterate until convergence (i.e., $\varepsilon_1 + \varepsilon_2$ small enough)
  (a) Find $\mathcal{A} = \mathrm{coalesce}(\mathcal{A}_1, \mathcal{A}_2)$ and run isotonic regression to minimize $f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$ such that w is compatible with $\mathcal{A}$.
  (b) Run accelerated Dykstra's algorithm to find the projection of 0 onto the intersection of $\widehat{B}^{\mathcal{A}_1}(F_1) - u/2 + w/2$ and $u/2 - w/2 - \widehat{B}^{\mathcal{A}_2}(F_2)$.
  (c) Merge the sets in $\mathcal{A}_j$ which are tight for s_j, $j \in \{1, 2\}$.
  (d) Check optimality by solving $\min_{C_{j,i_j} \subseteq A_{j,i_j}} F_j(B_{j,i_j-1} \cup C_{j,i_j}) - F_j(B_{j,i+j-1}) - s_j(C_{j,i_j})$ for $i_j \in \{1, \ldots, m_j\}$, Monitor ε_1 and ε_2 such that $F_j(C_j) - s_j(C_j) \geqslant -\varepsilon_j$, $j = 1, 2$.
  (e) If both s_1 and s_2 not optimal, for all C_{j,i_j} which are different from $\varnothing$ and A_{j,i_j}, split partitions.

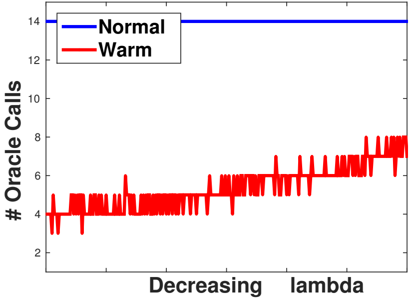
Active set method for decomposable functions

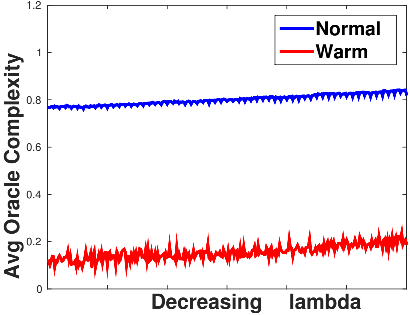
- **Input**: Submodular function F_1 and F_2 with SFM oracles, $u \in \mathbb{R}^n$, ordered partitions $\mathcal{A}_1, \mathcal{A}_2$
- **Algorithm**: iterate until convergence (i.e., $\varepsilon_1 + \varepsilon_2$ small enough)
  (a) Find $\mathcal{A} = \mathrm{coalesce}(\mathcal{A}_1, \mathcal{A}_2)$ and run isotonic regression to minimize $f(w) - u^\top w + \frac{1}{2}\|w\|_2^2$ such that w is compatible with $\mathcal{A}$.
  (b) Run accelerated Dykstra's algorithm to find the projection of 0 onto the intersection of $\widehat{B}^{\mathcal{A}_1}(F_1) - u/2 + w/2$ and $u/2 - w/2 - \widehat{B}^{\mathcal{A}_2}(F_2)$.
  (c) Merge the sets in $\mathcal{A}_j$ which are tight for s_j, $j \in \{1, 2\}$.
  (d) Check optimality by solving $\min_{C_{j,i_j} \subseteq A_{j,i_j}} F_j(B_{j,i_j-1} \cup C_{j,i_j}) - F_j(B_{j,i+j-1}) - s_j(C_{j,i_j})$ for $i_j \in \{1, \ldots, m_j\}$, Monitor ε_1 and ε_2 such that $F_j(C_j) - s_j(C_j) \geqslant -\varepsilon_j$, $j = 1, 2$.
  (e) If both s_1 and s_2 not optimal, for all C_{j,i_j} which are different from $\varnothing$ and A_{j,i_j}, split partitions.
- **Output**: $w \in \mathbb{R}^n$ and $s_1 \in B(F_1)$, $s_2 \in B(F_2)$.

Results for 2D TV without decomposition

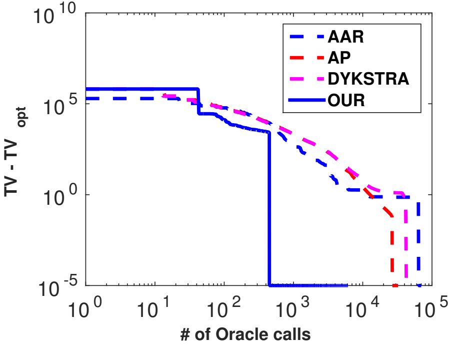


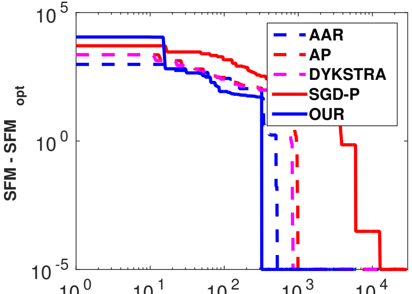
Results for 2D TV without decomposition (Warm Start)





Results for 2D TV with decomposition into 1D





Thank You. Questions?