On Polynomial-Time Solvability of Combinatorial Markov Random Fields

Shaoning Han

Daniel J. Epstein Department of Industrial & Systems Engineering
University of Southern California

shaoning@usc.edu

October, 2022

Collaborators

Andres Gomez
University of Southern California

Jong-Shi Pang
University of Southern California

Introduction to Markov random fields

Markov random field An MRF model is defined on an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where

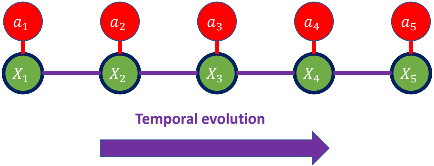
- random variable $X_i = x_i + \epsilon_i$ with $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ for $i \in \mathcal{V}$
- X_i is only dependent on its neighbors and independent of others

Introduction to Markov random fields

Markov random field An MRF model is defined on an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where

- random variable $X_i = x_i + \epsilon_i$ with $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ for $i \in \mathcal{V}$
- X_i is only dependent on its neighbors and independent of others

Applications

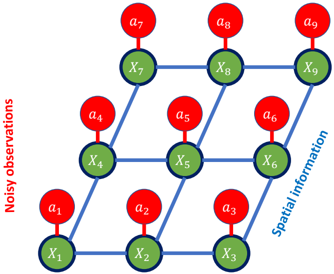


Introduction to Markov random fields

Markov random field An MRF model is defined on an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where

- random variable $X_i = x_i + \epsilon_i$ with $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ for $i \in \mathcal{V}$
- X_i is only dependent on its neighbors and independent of others

Applications

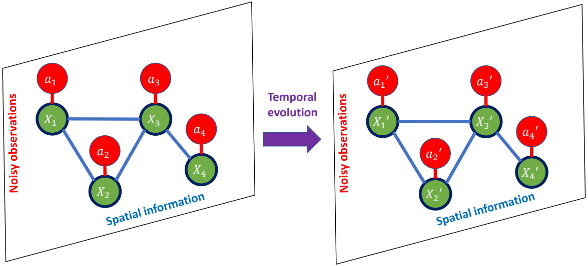


Introduction to Markov random fields

Markov random field An MRF model is defined on an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where

- random variable $X_i = x_i + \epsilon_i$ with $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ for $i \in \mathcal{V}$
- X_i is only dependent on its neighbors and independent of others

Applications



Markov random field An MRF model is defined on an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where

- random variable $X_i = x_i + \epsilon_i$ with $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ for $i \in \mathcal{V}$
- X_i is only dependent on its neighbors and independent of others

Applications

- signal processing (Hochbaum and Lu 2017)
- image denoising (Boykov and Funka-Lea 2006)
- epidemiology (Knorr-Held and Besag 1998)
- criminology (Law et al. 2014)
- bioinformatics (Eilers and De Menezes 2005)
- ...

Introduction to Markov random fields

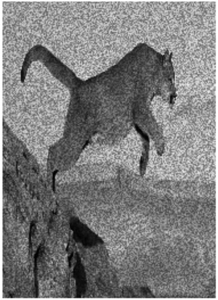
Given observations $\{a_i\}_{i\in\mathcal{V}}$, the MRF inference problem can be stated as

$$\begin{aligned}&\underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i)}_{\text{fitness}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{smoothness}}\\ &\textbf{subject to } \boldsymbol{\ell} \le \boldsymbol{x} \le \boldsymbol{u}\end{aligned}$$

- a_i: observations
- x_i: true values
- h_i, g_{ij}: convex, nonnegative with $h_i(0) = g_{ij}(0) = 0\ \forall\, i, j$
- $\ell_i \in \mathbb{R} \cup \{-\infty\}$, $u_i \in \mathbb{R} \cup \{+\infty\}$

Example - image denoising

Image denoising



Observation $\boldsymbol{a}$

Estimator $\boldsymbol{x}$

- x_i: value of pixels
- smoothness: adjacent pixels have similar values

Sparse MRF inference

Assumption the underlying statistical process is sparse

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i)}_{\text{fitness}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{smoothness}} + \underbrace{\lambda\|\boldsymbol{x}\|_0}_{\text{sparsity}}$$

subject to $\boldsymbol{\ell} \le \boldsymbol{x} \le \boldsymbol{u}$

- $\|\boldsymbol{x}\|_0$: the number of nonzero entries of $\boldsymbol{x}$

Sparse MRF inference

Assumption the underlying statistical process is sparse

$$\textbf{minimize}_{\boldsymbol{x}\in\mathbb{R}^n} \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i)}_{\text{fitness}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{smoothness}} + \underbrace{\lambda \sum_{i\in\mathcal{V}} z_i}_{\text{sparsity}}$$

$$\textbf{subject to } \boldsymbol{\ell} \circ \boldsymbol{z} \le \boldsymbol{x} \le \boldsymbol{u} \circ \boldsymbol{z}$$

$$\boldsymbol{z} \in \{0,1\}^n$$

- $\boldsymbol{u} \circ \boldsymbol{v} = (u_1 v_1, \ldots, u_n v_n)$ Hadamard product
- define $0 \cdot (\pm\infty) = 0$
- $[z_i = 0 \Rightarrow x_i = 0]$ & $[z_i = 1 \Rightarrow x_i \in [\ell_i, u_i]]$
- If $\ell_i = -\infty$ and $u_i = +\infty$, then $\ell_i z_i \le x_i \le u_i z_i \Leftrightarrow x_i(1 - z_i) = 0$

Robust MRF

Assumption a few observations a_i are corrupted by gross outliers

$$\underset{\mathcal{U}\subseteq\mathcal{V},\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \sum_{i\in\mathcal{V}\setminus\mathcal{U}} \underbrace{h_i(x_i - a_i)}_{\text{fitness}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{smoothness}} + \underbrace{\lambda|\mathcal{U}|}_{\text{robustness}}$$

$$\textbf{subject to } \ell \leq \boldsymbol{x} \leq \boldsymbol{u}$$

- $\mathcal{U}$: the set of outliers

Introducing binary variables $[z_i = 1 \Leftrightarrow i \in \mathcal{U}] \Rightarrow$ a MIP formulation

Robust MRF

Assumption a few observations a_i are corrupted by gross outliers

$$\min_{\boldsymbol{z},\boldsymbol{x}\in\mathbb{R}^n} \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i-a_i)(1-z_i)}_{\text{fitness}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i-x_j)}_{\text{smoothness}} + \underbrace{\lambda\sum_{i\in\mathcal{V}} z_i}_{\text{robustness}}$$

$$\text{subject to } \boldsymbol{\ell} \le \boldsymbol{x} \le \boldsymbol{u}$$

$$\boldsymbol{z}\in\{0,1\}^n$$

- $z_i = 1 \Leftrightarrow a_i$ is an outlier

Robust MRF

Assumption a few observations a_i are corrupted by gross outliers

$$\underset{\boldsymbol{z},\boldsymbol{w},\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i - w_i)}_{\text{fitness}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{smoothness}} + \underbrace{\lambda \sum_{i\in\mathcal{V}} z_i}_{\text{robustness}}$$

$$\textbf{subject to } \boldsymbol{\ell} \le \boldsymbol{x} \le \boldsymbol{u}$$

$$w_i(1 - z_i) = 0 \ \forall i \in \mathcal{V} \ , \ \boldsymbol{z} \in \{0,1\}^n$$

Equivalence

- $z_i = 0$: $w_i = 0 \Rightarrow a_i$ is not an outlier
- $z_i = 1$: $\min h_i(\bullet) = h_i(0) = 0 \Rightarrow w_i = a_i - x_i$ at the optimal solution $\Rightarrow$ a_i is an outlier

Robust MRF

Assumption a few observations a_i are corrupted by gross outliers

$$\underset{\boldsymbol{z},\boldsymbol{w},\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i - w_i)}_{\text{fitness}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{smoothness}} + \underbrace{\lambda \sum_{i\in\mathcal{V}} z_i}_{\text{robustness}}$$

$$\textbf{subject to } \boldsymbol{\ell} \le \boldsymbol{x} \le \boldsymbol{u}$$

$$\tilde{\boldsymbol{\ell}} \circ \boldsymbol{z} \le \boldsymbol{w} \le \tilde{\boldsymbol{u}} \circ \boldsymbol{z}, \;\; \boldsymbol{z} \in \{0,1\}^n$$

Equivalence

- $z_i = 0$: $w_i = 0 \Rightarrow a_i$ is not an outlier
- $z_i = 1$: $\min h_i(\bullet) = h_i(0) = 0 \Rightarrow w_i = a_i - x_i$ at the optimal solution $\Rightarrow$ a_i is an outlier
- $\tilde{\ell}_i = -\infty$ and $\tilde{u}_i = +\infty \;\; \forall\, i \in \mathcal{V}$

Main result

Consider a general combinatorial statistical inference problem

$$\underset{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ z \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\},$$

where $f(\bullet)$ is a convex function and $\boldsymbol{c} \geq 0$

- In general, the problem is $\mathcal{NP}$-hard, e.g., if $f(\bullet)$ = the obj of OLS

Main result

Consider a general combinatorial statistical inference problem

$$\underset{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ z \le \boldsymbol{x} \le \boldsymbol{u} \circ \boldsymbol{z} \right\},$$

where $f(\bullet)$ is a convex function and $\boldsymbol{c} \ge 0$

- In general, the problem is $\mathcal{NP}$-hard, e.g., if $f(\bullet)$ = the obj of OLS

However, in the context of MRF

Theorem (Han et al. (2022))

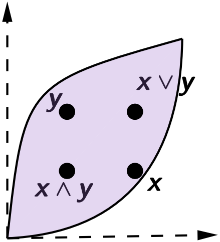
The problem of sparse/robust MRF inference can be solved as a binary submodular minimization problem and thus is (strongly) polynomially solvable.

Preliminaries - lattices

Meet and join Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, define

$$\text{meet} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \wedge \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \min\{x_1, y_1\} \\ \min\{x_2, y_2\} \\ \vdots \\ \min\{x_n, y_n\} \end{pmatrix}, \quad \text{join} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \vee \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \max\{x_1, y_1\} \\ \max\{x_2, y_2\} \\ \vdots \\ \max\{x_n, y_n\} \end{pmatrix}$$

Lattice A set $\mathcal{L} \subset \mathbb{R}^n$ is a lattice if $[\boldsymbol{x}, \boldsymbol{y} \in \mathcal{L} \Rightarrow \boldsymbol{x} \vee \boldsymbol{y}, \boldsymbol{x} \wedge \boldsymbol{y} \in \mathcal{L}]$

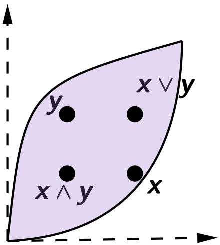


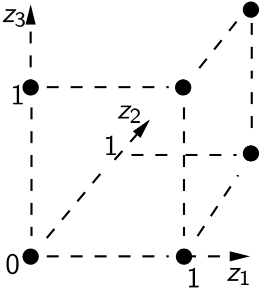
Preliminaries - lattices

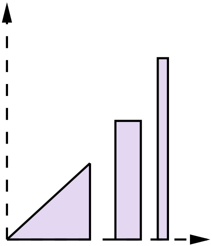
Meet and join Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, define

$$\text{meet} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \wedge \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \min\{x_1, y_1\} \\ \min\{x_2, y_2\} \\ \vdots \\ \min\{x_n, y_n\} \end{pmatrix}, \quad \text{join} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \vee \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \max\{x_1, y_1\} \\ \max\{x_2, y_2\} \\ \vdots \\ \max\{x_n, y_n\} \end{pmatrix}$$

Lattice A set $\mathcal{L} \subset \mathbb{R}^n$ is a lattice if $[\boldsymbol{x}, \boldsymbol{y} \in \mathcal{L} \Rightarrow \boldsymbol{x} \vee \boldsymbol{y}, \boldsymbol{x} \wedge \boldsymbol{y} \in \mathcal{L}]$







Preliminaries - submodularity

Submodularity Given a lattice $\mathcal{L} \subseteq \mathbb{R}^n$, a function $f : \mathcal{L} \to \mathbb{R}$ is submodular if

$$f(\boldsymbol{x}) + f(\boldsymbol{y}) \geq f(\boldsymbol{x} \wedge \boldsymbol{y}) + f(\boldsymbol{x} \vee \boldsymbol{y}) \ \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{L}$$

Remarks

- If $\mathcal{L} \subseteq \{0, 1\}^n$, then f is a binary/set submodular function
- If $f \in \mathcal{C}^2(\mathbb{R}^n)$, submodularity over $\mathbb{R}^n \Leftrightarrow \frac{\partial^2 f}{\partial y_i \partial y_j}(\boldsymbol{y}) \leq 0 \ \forall i \neq j$

Preliminaries - submodularity

Submodularity Given a lattice $\mathcal{L} \subseteq \mathbb{R}^n$, a function $f : \mathcal{L} \to \mathbb{R}$ is submodular if

$$f(\boldsymbol{x}) + f(\boldsymbol{y}) \geq f(\boldsymbol{x} \wedge \boldsymbol{y}) + f(\boldsymbol{x} \vee \boldsymbol{y}) \ \forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{L}$$

Remarks

- If $\mathcal{L} \subseteq \{0, 1\}^n$, then f is a binary/set submodular function
- If $f \in \mathcal{C}^2(\mathbb{R}^n)$, submodularity over $\mathbb{R}^n \Leftrightarrow \frac{\partial^2 f}{\partial y_i \partial y_j}(\boldsymbol{y}) \leq 0 \ \forall i \neq j$

Examples

- $n = 1 \Rightarrow f(x)$ is submodular
- $f(\boldsymbol{x}) = \boldsymbol{c}^\top \boldsymbol{x}$ is submodular
- $f(\boldsymbol{x}) = \boldsymbol{x}^\top \begin{bmatrix} 5 & -1 & -3 \\ -1 & 3 & -2 \\ -3 & -2 & 7 \end{bmatrix} \boldsymbol{x}$ is submodular

Two key lemmas

Lemma (Topkis (1978))

Given a lattice $\mathcal{L} \in \mathbb{R}^m \times \mathbb{R}^n$ *and a submodular function* $f : \mathcal{L} \to \mathbb{R}$, *the marginal function*

$$v(\boldsymbol{z}) \overset{def}{=} \underset{\boldsymbol{x} \in \mathbb{R}^m}{\textbf{minimum}} \{ f(\boldsymbol{x}, \boldsymbol{z}) : (\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{L} \}$$

is submodular on the lattice $\text{proj}_{\boldsymbol{z}} \overset{def}{=} \{ \boldsymbol{z} : \exists \boldsymbol{x} \text{ s.t. } (\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{L} \}$.

Two key lemmas

Lemma (Topkis (1978))

Given a lattice $\mathcal{L} \in \mathbb{R}^m \times \mathbb{R}^n$ and a submodular function $f : \mathcal{L} \to \mathbb{R}$, the marginal function

$$v(\boldsymbol{z}) \stackrel{def}{=} \underset{\boldsymbol{x} \in \mathbb{R}^m}{\textbf{minimum}} \{ f(\boldsymbol{x}, \boldsymbol{z}) : (\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{L} \}$$

is submodular on the lattice $\mathrm{proj}_{\boldsymbol{z}} \stackrel{def}{=} \{ \boldsymbol{z} : \exists \boldsymbol{x} \text{ s.t. } (\boldsymbol{x}, \boldsymbol{z}) \in \mathcal{L} \}$.

Lemma

If $g : \mathbb{R} \to \mathbb{R}$ is a convex function, then $\bar{g}(x, y) \stackrel{def}{=} g(x - y)$ is a submodular function over $\mathbb{R}^2$.

Nonnegative case

Assume $\boldsymbol{\ell} \geq \mathbf{0}$ and get back to

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\} \tag{$\star$}$$

Objective

$$f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} = \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i)}_{\text{univariate}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{Lemma}} + \underbrace{\boldsymbol{c}^\top \boldsymbol{z}}_{\text{linear}}$$

Feasible region $\prod_{i\in\mathcal{V}} \left\{ (x_i, z_i) \in \mathbb{R} \times \{0, 1\} : \ell_i z_i \leq x_i \leq z_i u_i \right\}$

Nonnegative case

Assume $\boldsymbol{\ell} \geq \mathbf{0}$ and get back to

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\} \tag{$\star$}$$

Objective is submodular

$$f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} = \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i)}_{\text{univariate}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{Lemma}} + \underbrace{\boldsymbol{c}^\top \boldsymbol{z}}_{\text{linear}}$$

Feasible region $\prod_{i\in\mathcal{V}} \left\{ (x_i, z_i) \in \mathbb{R} \times \{0,1\} : \ell_i z_i \leq x_i \leq z_i u_i \right\}$

Nonnegative case

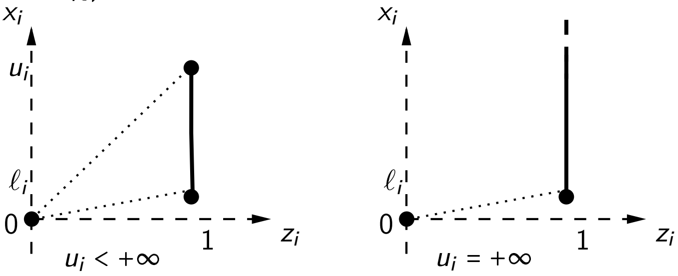
Assume $\boldsymbol{\ell} \geq \mathbf{0}$ and get back to

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\} \qquad (\star)$$

Objective is submodular

$$f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} = \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i)}_{\text{univariate}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{Lemma}} + \underbrace{\boldsymbol{c}^\top \boldsymbol{z}}_{\text{linear}}$$

Feasible region $\prod_{i\in\mathcal{V}} \left\{ (x_i, z_i) \in \mathbb{R} \times \{0,1\} : \ell_i z_i \leq x_i \leq z_i u_i \right\}$



Nonnegative case

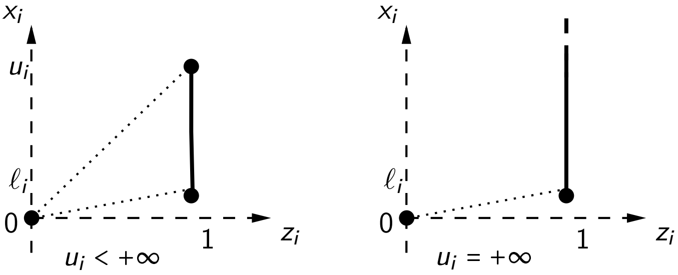
Assume $\ell \geq \mathbf{0}$ and get back to

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\} \tag{$\star$}$$

Objective is submodular

$$f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} = \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i)}_{\text{univariate}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{Lemma}} + \underbrace{\boldsymbol{c}^\top \boldsymbol{z}}_{\text{linear}}$$

Feasible region is a lattice



Nonnegative case

Assume $\boldsymbol{\ell} \geq \mathbf{0}$ and get back to

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\} \tag{$\star$}$$

Objective is submodular

$$f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} = \sum_{i\in\mathcal{V}} \underbrace{h_i(x_i - a_i)}_{\text{univariate}} + \sum_{(i,j)\in\mathcal{E}} \underbrace{g_{ij}(x_i - x_j)}_{\text{Lemma}} + \underbrace{\boldsymbol{c}^\top \boldsymbol{z}}_{\text{linear}}$$

Feasible region is a lattice

Thus,

$$(\star) \Leftrightarrow \underset{\boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}}\, v(\boldsymbol{z}) + \boldsymbol{c}^\top \boldsymbol{z}$$

where $v(\boldsymbol{z}) = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimum}} \left\{ f(\boldsymbol{x}) : \boldsymbol{\ell} \circ \boldsymbol{z} \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\}$ is a binary submodular function and can be evaluated by solving a convex program

Assume $\boldsymbol{\ell} \ngeq \mathbf{0}$,

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \le \boldsymbol{x} \le \boldsymbol{u} \circ \boldsymbol{z} \right\} \qquad (\star)$$

Issue the feasible region is not a lattice if $\ell_i < 0$

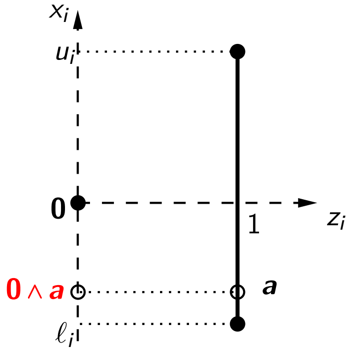


Figure: Region of $\{(x_i, z_i) \in \mathbb{R} \times \{0,1\} : \ell_i z_i \le x_i \le z_i u_i\}$

General case

Assume $\ell \ngeq \mathbf{0}$,

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \le \boldsymbol{x} \le \boldsymbol{u} \circ \boldsymbol{z} \right\} \qquad (\star)$$

Issue the feasible region is not a lattice if $\ell_i < 0$

Idea If $\ell_i < 0$ and $u_i > 0$, then

$$\ell_i z_i \le x_i \le u_i z_i,\ z_i \in \{0,1\} \Leftrightarrow \begin{cases} z_i = z_i^+ + (1 - z_i^-) \\ z_i^+ \in \{0,1\},\ z_i^- \in \{0,1\} \\ \ell_i (1 - z_i^-) \le z_i \le u_i z_i^+ \\ z_i^- \ge z_i^+ \end{cases}$$

- Split z_i into two parts z_i^+ and $(1 - z_i^-)$

General case

Assume $\boldsymbol{\ell} \ngeq \mathbf{0}$,

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\} \qquad (\star)$$

Issue the feasible region is not a lattice if $\ell_i < 0$

Idea If $\ell_i < 0$ and $u_i > 0$, then

$$\ell_i z_i \leq x_i \leq u_i z_i, \; z_i \in \{0,1\} \Leftrightarrow \begin{cases} z_i = z_i^+ + (1 - z_i^-) \\ z_i^+ \in \{0,1\}, \; z_i^- \in \{0,1\} \\ \ell_i(1 - z_i^-) \leq z_i \leq u_i z_i^+ \\ z_i^- \geq z_i^+ \end{cases}$$

- Split z_i into two parts z_i^+ and $(1 - z_i^-)$
- $z_i^+ = 0 \Rightarrow [x_i]_+ \overset{\text{def}}{=} \max\{x_i, 0\} = 0 \Leftrightarrow x_i \leq 0$

General case

Assume $\boldsymbol{\ell} \ngeq \mathbf{0}$,

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \le \boldsymbol{x} \le \boldsymbol{u} \circ \boldsymbol{z} \right\} \qquad (\star)$$

Issue the feasible region is not a lattice if $\ell_i < 0$

Idea If $\ell_i < 0$ and $u_i > 0$, then

$$\ell_i z_i \le x_i \le u_i z_i,\ z_i \in \{0,1\} \Leftrightarrow \begin{cases} z_i = z_i^+ + (1 - z_i^-) \\ z_i^+ \in \{0,1\},\ z_i^- \in \{0,1\} \\ \ell_i(1 - z_i^-) \le z_i \le u_i z_i^+ \\ z_i^- \ge z_i^+ \end{cases}$$

- Split z_i into two parts z_i^+ and $(1 - z_i^-)$
- $z_i^+ = 0 \Rightarrow [x_i]_+ \overset{\text{def}}{=} \max\{x_i, 0\} = 0 \Leftrightarrow x_i \le 0$
- $1 - z_i^- = 0 \Rightarrow [x_i]_- \overset{\text{def}}{=} \max\{-x_i, 0\} = 0 \Leftrightarrow x_i \ge 0$

General case

Assume $\boldsymbol{\ell} \ngeq \mathbf{0}$,

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z} \right\} \qquad (\star)$$

Issue the feasible region is not a lattice if $\ell_i < 0$

Idea If $\ell_i < 0$ and $u_i > 0$, then

$$\ell_i z_i \leq x_i \leq u_i z_i,\ z_i \in \{0,1\} \Leftrightarrow \begin{cases} z_i = z_i^+ + (1 - z_i^-) \\ z_i^+ \in \{0,1\},\ z_i^- \in \{0,1\} \\ \ell_i(1 - z_i^-) \leq z_i \leq u_i z_i^+ \\ z_i^- \geq z_i^+ \end{cases}$$

- Split z_i into two parts z_i^+ and $(1 - z_i^-)$
- $z_i^+ = 0 \Rightarrow [x_i]_+ \stackrel{\text{def}}{=} \max\{x_i, 0\} = 0 \Leftrightarrow x_i \leq 0$
- $1 - z_i^- = 0 \Rightarrow [x_i]_- \stackrel{\text{def}}{=} \max\{-x_i, 0\} = 0 \Leftrightarrow x_i \geq 0$
- $[x_i]_+$ and $[x_i]_-$ can not be both nonzero $\Rightarrow z_i^+ + (1 - z_i^-) \leq 1$

General case

Assume $\ell \ngeq \mathbf{0}$,

$$\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{z}\in\{0,1\}^n}{\textbf{minimize}} \left\{ f(\boldsymbol{x}) + \boldsymbol{c}^\top \boldsymbol{z} : \boldsymbol{\ell} \circ \boldsymbol{z} \le \boldsymbol{x} \le \boldsymbol{u} \circ \boldsymbol{z} \right\} \qquad (\star)$$

Issue the feasible region is not a lattice if $\ell_i < 0$

Idea If $\ell_i < 0$ and $u_i > 0$, then

$$\ell_i z_i \le x_i \le u_i z_i,\ z_i \in \{0,1\} \Leftrightarrow \begin{cases} z_i = z_i^+ + (1 - z_i^-) \\ z_i^+ \in \{0,1\},\ z_i^- \in \{0,1\} \\ \ell_i(1 - z_i^-) \le z_i \le u_i z_i^+ \\ z_i^- \ge z_i^+ \end{cases} \Rightarrow \text{lattice}$$

- Split z_i into two parts z_i^+ and $(1 - z_i^-)$
- $z_i^+ = 0 \Rightarrow [x_i]_+ \overset{\text{def}}{=} \max\{x_i, 0\} = 0 \Leftrightarrow x_i \le 0$
- $1 - z_i^- = 0 \Rightarrow [x_i]_- \overset{\text{def}}{=} \max\{-x_i, 0\} = 0 \Leftrightarrow x_i \ge 0$
- $[x_i]_+$ and $[x_i]_-$ can not be both nonzero $\Rightarrow z_i^+ + (1 - z_i^-) \le 1$

General case

For simplicity, assume $\boldsymbol{\ell} < \mathbf{0} < \boldsymbol{u}$. Substituting out z_i,

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{z}^+,\boldsymbol{z}^- \in \mathbb{R}^n}{\textbf{minimize}}\ & f(\boldsymbol{x}) + \boldsymbol{c}^\top(\boldsymbol{z}^+ + \mathbf{1} - \boldsymbol{z}^-) \\ \textbf{subject to}\ & \boldsymbol{\ell} \circ (\mathbf{1} - \boldsymbol{z}^-) \le \boldsymbol{x} \le \boldsymbol{u} \circ \boldsymbol{z}^+ \\ & \boldsymbol{z}^- \ge \boldsymbol{z}^+,\ \boldsymbol{z}^+, \boldsymbol{z}^- \in \{0,1\}^n \end{aligned} \tag{$\star$}$$

General case

For simplicity, assume $\boldsymbol{\ell} < \mathbf{0} < \boldsymbol{u}$. Substituting out z_i,

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{z}^+,\boldsymbol{z}^- \in \mathbb{R}^n}{\textbf{minimize}} \ & f(\boldsymbol{x}) + \boldsymbol{c}^\top(\boldsymbol{z}^+ + \mathbf{1} - \boldsymbol{z}^-) \\ \textbf{subject to} \ & \boldsymbol{\ell} \circ (\mathbf{1} - \boldsymbol{z}^-) \le \boldsymbol{x} \le \boldsymbol{u} \circ \boldsymbol{z}^+ \\ & \cancel{\boldsymbol{z}^- \ge \boldsymbol{z}^+},\ \boldsymbol{z}^+, \boldsymbol{z}^- \in \{0,1\}^n \end{aligned} \tag{$\star$}$$

General case

For simplicity, assume $\boldsymbol{\ell} < \mathbf{0} < \boldsymbol{u}$. Substituting out z_i,

$$\begin{aligned} \underset{\boldsymbol{x}, \boldsymbol{z}^+, \boldsymbol{z}^- \in \mathbb{R}^n}{\textbf{minimize}} \; & f(\boldsymbol{x}) + \boldsymbol{c}^\top (\boldsymbol{z}^+ + \mathbf{1} - \boldsymbol{z}^-) \\ \textbf{subject to} \; & \boldsymbol{\ell} \circ (\mathbf{1} - \boldsymbol{z}^-) \leq \boldsymbol{x} \leq \boldsymbol{u} \circ \boldsymbol{z}^+ \\ & \cancel{\boldsymbol{z}^- \geq \boldsymbol{z}^+}, \; \boldsymbol{z}^+, \boldsymbol{z}^- \in \{0,1\}^n \end{aligned} \tag{$\star$}$$

A mixed-integer submodular minimization problem! $\Rightarrow$

General case

For simplicity, assume $\boldsymbol{\ell} < \mathbf{0} < \boldsymbol{u}$. Substituting out z_i,

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{z}^+,\boldsymbol{z}^-\in\mathbb{R}^n}{\textbf{minimize}}\ & f(\boldsymbol{x}) + \boldsymbol{c}^\top(\boldsymbol{z}^+ + \mathbf{1} - \boldsymbol{z}^-) \\ \textbf{subject to}\ & \boldsymbol{\ell}\circ(\mathbf{1}-\boldsymbol{z}^-) \le \boldsymbol{x} \le \boldsymbol{u}\circ\boldsymbol{z}^+ \\ & \cancel{\boldsymbol{z}^- \ge \boldsymbol{z}^+},\ \boldsymbol{z}^+,\boldsymbol{z}^- \in \{0,1\}^n \end{aligned} \tag{$\star$}$$

A mixed-integer submodular minimization problem! $\Rightarrow$

$$(\star) \Leftrightarrow \underset{(\boldsymbol{z}^+,\boldsymbol{z}^-)\in\{0,1\}^{2n}}{\textbf{minimize}}\ v(\boldsymbol{z}^+,\boldsymbol{z}^-) + \boldsymbol{c}^\top(\boldsymbol{z}^+ + \mathbf{1} - \boldsymbol{z}^-),$$

where $v(\boldsymbol{z}^+,\boldsymbol{z}^-) \overset{\text{def}}{=} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimum}} \{f(\boldsymbol{x}) : \boldsymbol{\ell}\circ(\mathbf{1}-\boldsymbol{z}^-) \le \boldsymbol{x} \le \boldsymbol{u}\circ\boldsymbol{z}^+\}$ is a binary submodular function and can be evaluated by solving a convex program

Take home message

- Sparse/robust MRF inference problems are polynomially solvable!

Our paper is available at: https://arxiv.org/abs/2209.13161

Thank You!

Reference I

Boykov, Y. and Funka-Lea, G. (2006). Graph cuts and efficient ND image segmentation. International Journal of Computer Vision, 70(2):109–131.

Eilers, P. H. and De Menezes, R. X. (2005). Quantile smoothing of array CGH data. Bioinformatics, 21(7):1146–1153.

Han, S., Pang, J.-S., and Gómez, A. (2022). On polynomial-time solvability of combinatorial markov random fields. https://arxiv.org/abs/2209.13161.

Hochbaum, D. S. and Lu, C. (2017). A faster algorithm solving a generalization of isotonic median regression and a class of fused lasso problems. SIAM Journal on Optimization, 27(4):2563–2596.

Knorr-Held, L. and Besag, J. (1998). Modelling risk from a disease in time and space. Statistics in Sedicine, 17(18):2045–2060.

Law, J., Quick, M., and Chan, P. (2014). Bayesian spatio-temporal modeling for analysing local patterns of crime over time at the small-area level. Journal of Quantitative Criminology, 30(1):57–78.

Reference II

Topkis, D. M. (1978). Minimizing a submodular function on a lattice. Operations research, 26(2):305–321.