

Flip-Flop Sublinear Models for Graphs

Supplementary Material

Proof of Theorem 1

Brijnesh Jain

Technische Universität Berlin, Germany
brijnesh.jain@gmail.com

Abstract. We prove that there is no class-dual for almost all sublinear models on graphs.

1 Introduction

We prove the following result:

Theorem 1. *There is no class-dual of a sublinear function with probability one.*

The proof is based on material presented in

[1] B. Jain. *Flip-Flop Sublinear Models for Graphs*, S+SSPR 2014.

The content of [1] is not included in this contribution. Section 1 introduces the formalism necessary to derive the proof. Section 2 proves auxiliary results. Finally, the proof of Theorem 1 is presented in Section 3.

2 Preliminaries

Let \mathcal{G} be a permutation group and $\mathcal{X}_{\mathcal{G}}$ be the quotient of the group action \mathcal{G} on the matrix space $\mathcal{X} = \mathbb{R}^{n \times n}$, where n is some number not less than the number of vertices of the largest graphs. By π we denote the natural projection from \mathcal{X} to $\mathcal{X}_{\mathcal{G}}$. For details we refer to [1,2]. The stabilizer of $\mathbf{w} \in \mathcal{X}$ is a subgroup defined by

$$\mathcal{G}_{\mathbf{w}} = \{\gamma \in \mathcal{G} : \gamma \mathbf{w} = \mathbf{w}\}.$$

If $\mathcal{G}_{\mathbf{w}}$ is the trivial group, then each element of the orbit $[\mathbf{w}]$ has a trivial stabilizer. A graph X is said to be regular, if there is a representation $\mathbf{x} \in X$ with trivial stabilizer $\mathcal{G}_{\mathbf{x}}$.

Suppose that $f(X) = W \cdot X + b$ is a sublinear function. A cross section with basepoint $\mathbf{w} \in \mathcal{X}$ is an injective map $\phi : \mathcal{X}_{\mathcal{G}} \rightarrow \mathcal{X}$ satisfying

1. $W \cdot X = \phi(W)^T \phi(X)$
2. $\pi \circ \phi(X) = X$

for all $X \in \mathcal{X}_G$. We may regard the map ϕ as an isometric embedding of the graph space into some Euclidean space with respect to $\mathbf{w} = \phi(W)$. Note that ϕ is not uniquely determined, even for fixed $\mathbf{w} = \phi(W)$. The closure of the image $\phi(\mathcal{X}_G)$ is the Dirichlet (fundamental) domain with basepoint (center) \mathbf{w} defined by

$$\mathcal{D}_{\mathbf{w}} = \left\{ \mathbf{x} \in \mathcal{X} : \mathbf{w}^T \mathbf{x} \geq \tilde{\mathbf{w}}^T \mathbf{x}, \tilde{\mathbf{w}} \in W \right\}.$$

A Dirichlet domain is a convex polyhedral cone of dimension $\dim(\mathcal{X})$ with the following properties: (1) $\mathcal{D}_{\mathbf{w}}$ is well-defined, (2) $\mathbf{x} \in \mathcal{D}_{\mathbf{w}}$ iff (\mathbf{x}, \mathbf{w}) is an optimal alignment, (3) $\pi(\mathcal{D}_{\mathbf{w}}) = \mathcal{X}_G$, and (4) π is injective on the interior of $\mathcal{D}_{\mathbf{w}}$ [2].

3 Auxiliary Results

Lemma 1. *Let $\phi : \mathcal{X}_G \rightarrow \mathcal{X}$ be a fundamental cross section with basepoint \mathbf{w} . Then $\gamma \circ \phi$ is a fundamental cross section with basepoint $\gamma\mathbf{w}$ for all $\gamma \in \mathcal{G}$.*

Proof. From the axioms of a group action follows that γ is bijective. A cross section ϕ is injective by definition. Then the composition $\phi' = \gamma \circ \phi$ is also injective. We show that $W \cdot X = \phi'(W)^T \phi'(X)$. We have

$$\phi'(W)^T \phi'(X) = \gamma \mathbf{w}^T \gamma \mathbf{x},$$

where $\mathbf{w} = \phi(W)$ and $\mathbf{x} = \phi(X)$. Since \mathcal{G} is a permutation group, the mapping γ is orthogonal. As an orthogonal mapping, γ preserves the inner product, that is $\gamma \mathbf{w}^T \gamma \mathbf{x} = \mathbf{w}^T \mathbf{x}$. Since ϕ is a fundamental cross section with basepoint \mathbf{w} , we have

$$W \cdot X = \phi(W)^T \phi(X) = \mathbf{w}^T \mathbf{x} = \gamma \mathbf{w}^T \gamma \mathbf{x} = \phi'(W)^T \phi'(X).$$

Note that this part also shows that the ϕ' is a cross section with basepoint $\gamma\mathbf{w}$. Finally, we show that $\pi \circ \phi'(X) = X$ for all $X \in \mathcal{X}_G$. Since ϕ is a fundamental cross section, the vector $\mathbf{x} = \phi(X)$ projects to X via π . The graph X can be regarded as the orbit $[\mathbf{x}]$ of \mathbf{x} . As an element of $[\mathbf{x}]$ the vector $\gamma \mathbf{x}$ also projects to X via π . This shows $\pi \circ \phi'(X) = X$. \square

Lemma 2. *Let $\phi : \mathcal{X}_G \rightarrow \mathcal{X}$ be a fundamental cross section with basepoint \mathbf{w} and Dirichlet domain $\mathcal{D}_{\mathbf{w}}$. Then \mathbf{w} is an interior point of $\mathcal{D}_{\mathbf{w}}$ if and only if $W = \pi(\mathbf{w})$ is regular.*

Proof. Let \mathbf{w} be an interior point of $\mathcal{D}_{\mathbf{w}}$. We assume that W is not regular. Then there is a $\gamma \in \mathcal{G} \setminus \{\text{id}\}$ such that $\mathbf{w} = \gamma\mathbf{w}$. According to Lemma 1, the composition $\phi' = \gamma \circ \phi$ is a fundamental cross section with basepoint $\gamma\mathbf{w}$ and Dirichlet domain $\gamma\mathcal{D}_{\mathbf{w}} = \mathcal{D}_{\gamma\mathbf{w}}$. From $\mathbf{w} = \gamma\mathbf{w}$ follows that

$$\mathcal{D} = \text{int } \mathcal{D}_{\mathbf{w}} \cap \text{int } \gamma\mathcal{D}_{\mathbf{w}} \neq \emptyset,$$

where $\text{int } \mathcal{S}$ denotes the interior of a set $\mathcal{S} \subseteq \mathcal{X}$. As an intersection of open convex sets, the set \mathcal{D} is also open and convex. Then there are $n+1$ points $\mathbf{x}_0, \dots, \mathbf{x}_n \in \mathcal{D}$

in general position, where $n = \dim(\mathcal{X})$. By definition of a Dirichlet domain, we have

$$\mathbf{w}^T \mathbf{x}_i = \mathbf{w}^T \gamma \mathbf{x}_i$$

for all $i \in \{0, \dots, n\}$. This implies $\mathbf{x}_i = \gamma \mathbf{x}_i$, because cross sections are injective. Observe that \mathcal{G} acts isometrically on \mathcal{X} . Since two isometries are the same if they coincide on $n + 1$ points in general position, we obtain $\gamma = \text{id}$. This contradicts our assumption that there is a $\gamma \neq \text{id}$ such that $\mathbf{w} = \gamma \mathbf{w}$. Hence, W is regular.

Now we assume that W is regular and show that there is a representation \mathbf{w} of W such that $\mathbf{w} \in \text{int } \mathcal{D}_{\mathbf{w}}$. The boundary of $\mathcal{D}_{\mathbf{w}}$ is of the form

$$\text{bd } \mathcal{D}_{\mathbf{w}} = \bigcup_{\gamma \in \mathcal{G} \setminus \{\text{id}\}} \mathcal{D}_{\mathbf{w}} \cap \gamma \mathcal{D}_{\mathbf{w}}.$$

Since W is regular, $\gamma \mathbf{w} \neq \mathbf{w}$ for all $\gamma \in \mathcal{G} \setminus \{\text{id}\}$

Suppose that $\tilde{\mathbf{w}} \neq \mathbf{w}$ is another representation of W with Dirichlet domain $\mathcal{D}_{\tilde{\mathbf{w}}}$. The bisection of $\mathcal{D}_{\mathbf{w}}$ and $\mathcal{D}_{\tilde{\mathbf{w}}}$ is defined by the set

$$\mathcal{H}(\mathbf{w}, \tilde{\mathbf{w}}) = \{\mathbf{x} \in \mathcal{D} : \mathbf{w}^T \mathbf{x} = \tilde{\mathbf{w}}^T \mathbf{x}\},$$

where $\mathcal{D} = \mathcal{D}_{\mathbf{w}} \cap \mathcal{D}_{\tilde{\mathbf{w}}}$. Since \mathbf{w} and $\tilde{\mathbf{w}}$ are unequal, $\mathcal{H}(\mathbf{w}, \tilde{\mathbf{w}})$ is a subset of a hyperplane \mathcal{H} defined by the equation $h(\mathbf{x}) = (\mathbf{w} - \tilde{\mathbf{w}})^T \mathbf{x}$. The hyperplane \mathcal{H} is perpendicular to the vector $\mathbf{v} = \mathbf{w} - \tilde{\mathbf{w}}$ and passes through the midpoint of the connecting line between \mathbf{w} and $\tilde{\mathbf{w}}$. This shows that \mathbf{w} is not a point on $\mathcal{H}(\mathbf{w}, \tilde{\mathbf{w}})$. Since $\tilde{\mathbf{w}}$ was chosen arbitrarily, \mathbf{w} is in the interior of $\mathcal{D}_{\mathbf{w}}$. \square

Lemma 3. *Let $W \in \mathcal{X}_{\mathcal{G}}$ be regular. Suppose that ϕ is a fundamental cross section with basepoint \mathbf{w} and Dirichlet domain $\mathcal{D}_{\mathbf{w}}$. Then $-\mathbf{w} \notin \mathcal{D}_{\mathbf{w}}$.*

Proof. Observe that $\mathbf{w} \neq \mathbf{0}$ and $\mathbf{w} \neq -\mathbf{w}$ for all representations \mathbf{w} of a regular graph W . In addition, with W the graph $W' = \pi(-\mathbf{w})$ is also regular. As shown in [1], we have

$$W \cdot W = \max_{\gamma \in \mathcal{G}} \gamma \mathbf{w}^T \mathbf{w} = \mathbf{w}^T \mathbf{w}. \quad (1)$$

Let W' be the graph represented by $-\mathbf{w}$. From eq. (1) follows

$$-\mathbf{w}^T \mathbf{w} = -\min_{\gamma} \gamma \mathbf{w}^T \mathbf{w} < \max_{\gamma} -\mathbf{w}^T \mathbf{w} = W' \cdot W \quad (2)$$

Strict inequality in eq. (2) follows from regularity of W and W' together with $\mathbf{w} \neq -\mathbf{w}$. Thus, $-\mathbf{w}$ is not an element of the Dirichlet domain $\mathcal{D}_{\mathbf{w}}$ with basepoint \mathbf{w} . \square

Proof of Theorem 1

Suppose that each graph $X \in \mathcal{X}_G$ has a class label $y \in \mathcal{Y} = \{\pm 1\}$. By $\mathbb{C}[f]$ we denote the expected misclassification error of the sublinear function f . Consider a sublinear function of the form $f(X) = W \cdot X + b$, where W is regular and $b \neq 0$. The equation $f(X) = 0$ defines a decision surface \mathcal{H}_f that separates the graph space \mathcal{X}_G into two disjoint regions $\mathcal{R}_+(f)$ and $\mathcal{R}_-(f)$. By construction we have $\mathbb{C}[f] = 0$.

Let ϕ be a fundamental cross section with basepoint $\mathbf{w} = \phi(W)$ and Dirichlet domain $\mathcal{D}_{\mathbf{w}}$. By $\mathcal{D}_+(f) = \phi(\mathcal{R}_+(f))$ and $\mathcal{D}_-(f) = \phi(\mathcal{R}_-(f))$ we denote the images of both class regions $\mathcal{R}_+(f)$ and $\mathcal{R}_-(f)$ in $\mathcal{D}_{\mathbf{w}}$. The hyperplane separating $\mathcal{D}_+(f)$ and $\mathcal{D}_-(f)$ is defined by the equation $h(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b = 0$. By construction, the expected misclassification error of $h(\mathbf{x})$ is $\mathbb{C}_{\mathcal{X}}[h] = 0$. In addition, $h(\mathbf{x})$ is the unique global minimum of $\mathbb{C}_{\mathcal{X}}[\cdot]$ over all linear functions.

Now we relabel both class regions $\mathcal{R}_+(f)$ and $\mathcal{R}_-(f)$ such that all graphs from the positive class region $\mathcal{R}_+(f)$ have negative labels and all graphs from the negative class region $\mathcal{R}_-(f)$ have positive labels. We denote the relabeled regions in \mathcal{X}_G by $\bar{\mathcal{R}}_+(f)$ and $\bar{\mathcal{R}}_-(f)$, resp., and similarly the relabeled regions in $\mathcal{D}_{\mathbf{w}}$ by $\bar{\mathcal{D}}_+(f)$ and $\bar{\mathcal{D}}_-(f)$. For the relabeled variant, the linear classifier in $\mathcal{D}_{\mathbf{w}}$ determined by $h'(\mathbf{x}) = -\mathbf{w}^T \mathbf{x} + b$ has also minimal misclassification error $\mathbb{C}_{\mathcal{X}}[h'] = 0$ and is the unique minimizer of $\mathbb{C}_{\mathcal{X}}[\cdot]$. By Lemma 3 the opposite direction $-\mathbf{w}$ of basepoint \mathbf{w} is not an element of the Dirichlet domain $\mathcal{D}_{\mathbf{w}}$.

Suppose that $f'(X) = W' \cdot X + b'$ is a sublinear function such that

$$\mathbb{C}[f'] = \min_g \mathbb{C}[g]$$

under the relabeled setting with class regions $\bar{\mathcal{R}}_+(f)$ and $\bar{\mathcal{R}}_-(f)$. Let $\mathbf{w}' = \phi(W')$ be a representation of W' in $\mathcal{D}_{\mathbf{w}}$ and let ϕ' be a fundamental cross section with basepoint $\mathbf{w}' = \phi'(W')$ and Dirichlet domain $\mathcal{D}_{\mathbf{w}'}$. We consider the intersection $\mathcal{D} = \mathcal{D}_{\mathbf{w}} \cap \mathcal{D}_{\mathbf{w}'}$. According to [3], Theorem 12, the intersection of two convex polyhedral cones is again a convex polyhedral cone. We show that \mathcal{D} has co-dimension 0. For this, we first show that the relative interior of the intersection \mathcal{D} contains an inner point \mathbf{z} from $\mathcal{D}_{\mathbf{w}}$. Since W is regular, \mathbf{w} is an inner point of $\mathcal{D}_{\mathbf{w}}$ lying in \mathcal{D} . Two cases can occur: (i) \mathbf{w} is in the relative interior of \mathcal{D} ; and (ii) \mathbf{w} lies on the boundary of \mathcal{D} . If (i) holds, we set $\mathbf{z} = \mathbf{w}$. Suppose that (ii) holds. Since \mathbf{w} and \mathbf{w}' are distinct elements of the intersection \mathcal{D} , any point $\mathbf{z} = \lambda \mathbf{w} + (1 - \lambda) \mathbf{w}'$ with $0 < \lambda < 1$ lies in the relative interior of \mathcal{D} by convexity.

Next we show that there is an $\varepsilon > 0$ such that the open ball $\mathcal{B}(\mathbf{z}, \varepsilon)$ with center \mathbf{z} and radius ε is contained in \mathcal{D} . We choose $\varepsilon < \min(\lambda, 1 - \lambda)$. Then $\mathcal{B}(\mathbf{z}, \varepsilon) \subset \mathcal{D}_{\mathbf{w}}$, because \mathbf{z} is an interior point of $\mathcal{D}_{\mathbf{w}}$. Suppose that $\mathcal{B}(\mathbf{z}, \varepsilon)$ is not contained in \mathcal{D} . Since \mathbf{z} lies in the relative interior of \mathcal{D} , the co-dimension of \mathcal{D} is positive. In addition, \mathbf{z} lies on the boundary of $\mathcal{D}_{\mathbf{w}'}$. This implies that \mathbf{z} is also a boundary point of $\mathcal{D}_{\mathbf{w}}$, which is a contradiction of our construction. Hence, the co-dimension of \mathcal{D} is zero.

Since $b \neq 0$, the hyperplane defined by equation $h(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b = 0$ does not pass through the origin $\mathbf{0}$. Boundaries of any Dirichlet domain are

supported by hyperplanes passing through $\mathbf{0}$. Thus the intersection \mathcal{D} contains an open set \mathcal{U} separated by the hyperplane segment \mathcal{H}_f . Then there are $n + 1$ points $\mathbf{x}_0, \dots, \mathbf{x}_n \in \mathcal{D}$ in general position that fix the unique hyperplane. Since $-\mathbf{w} \notin \mathcal{D}_{\mathbf{w}}$, we have $\mathbb{C}[f'] > 0$ implying that f' is not a dual of f .

The probabilistic statement follows from the fact that interior points are regular and therefore have Lebesgue measure one. With respect to the bias b , the set $\mathbb{R} \setminus \{0\}$ also has Lebesgue measure one. Combining all parts we arrive at the assertion. \square

References

1. B. Jain and K. Obermayer. Structure Spaces. *The Journal of Machine Learning Research*, 10:2667–2714, 2009.
2. B. Jain and K. Obermayer. Learning in Riemannian Orbifolds. *arXiv preprint arXiv:1204.4294*, 2012.
3. M. Gerstenhaber. Theory of convex polyhedral cones. *Activity analysis of production and allocation*, 298–316, 1951.