
Analyzing PlanarCC: Demonstrating the Equivalence of PlanarCC and The Multi-Cut LP Relaxation

Julian Yarkony
Experian Data Lab
San Diego CA, 92075
julian.yarkony@experian.com

Abstract

Correlation clustering is an exciting area of research in the fields of graphical models and image segmentation. In this article we study the linear programming (LP) relaxation corresponding to PlanarCC and the multi-cut LP relaxation which are two methods for correlation clustering. We demonstrate that they have equal value when optimized. This provides justification for the use of PlanarCC which is very fast on planar graphs in situations where the multi-cut LP relaxation is currently employed.

1 Introduction

Correlation clustering Bansal et al. (2002) is a mathematically interesting area at the intersection of the domains of graphical models and image segmentation. In the domain of image segmentation, correlation clustering is used to indicate the boundaries between the regions of an image.

Correlation clustering for image segmentation relies on a local classifier such as the global probability of boundary, *gPb* Martin et al. (2004) to produce a probability that any pair of adjacent pixels are in separate regions. The prediction is based on local image features. The goal is to partition the image into an arbitrary number of regions so as to respect these local probabilities.

More generally super-pixels are used instead of pixels. A super-pixel is simply a small compact group of pixels in the same area of the image that have similar color values. Each super-pixel is closed meaning that there exists a path on the image between each pair of pixels in a super-pixel such that this path does not include pixels outside of the super-pixel. Super-pixels are non-overlapping meaning that they don't share pixels with each other. Super-pixels are constructed so as to ensure that it is unlikely that any given super-pixel is on both sides of a boundary in the image. Super-pixels are used to aggregate information so as to avoid hallucinating boundaries.

Correlation clustering in the domain of image segmentation has been approached using multiple methods such as the multi-cut (MC) linear programming (LP) relaxation or its integer programming version Andres et al. (2011), graph cut methods Bagon and Galun (2011), or PlanarCC Yarkony et al. (2012) and its extension amongst others. Correlation clustering is NP hard even on planar graphs Bachrach et al. (2011). One interesting property of correlation clustering in the domain of images is that the MC LP relaxation, and the PlanarCC LP relaxation appear to be generally tight and when not tight they are nearly tight Yarkony et al. (2012). When they are not tight naive rounding schemes produce solutions that are either globally optimal or very near global optimal.

PlanarCC is a LP relaxation based method that leverages the fact that *gPb* only provides probabilities between adjacent super-pixels thus creating a planar correlation clustering instance. The application of PlanarCC has been demonstrated to yield large increases in speed over the other LP based methods

⁰I would like to thank David Sontag for proposing this problem. I would also like to thank my managers at Experian for their continued support of my academic research and desire to publish my research.

Yarkony et al. (2012). In this paper we study the LP corresponding to PlanarCC in the context of the MC LP relaxation. We demonstrate that they correspond to the same LP relaxation. In this paper we first introduce the MC and take its dual, then we then introduce PlanarCC and take its dual. Then we show that the optimal value of the PlanarCC dual is upper and lower bounded by the optimal value of the MC dual meaning that they are equal and thus their corresponding primal problems also have equal optimizing objective value.

2 Correlation Clustering and the Multi-Cut (MC) LP Relaxation

Correlation clustering is formulated as follows. Consider a graph G where nodes refer to super-pixels and edges indicate adjacency. We denote a partition (also called a clustering or a segmentation in the literature) using binary indicator X . Here X is indexed by e where $X_e = 1$ if and only if a boundary is created on edge e . Here each edge e is associated with a real valued cost θ_e . Here θ_e refers to the cost to ‘cut’ (meaning create a boundary) at the edge e . Negative values of θ_e are associated with inclination to create a boundary at edge e and positive values are associated with an inclination to not create a boundary at edge e .

The objective of correlation clustering is to find the partition X that minimizes the sum of the cut edges subject to the constraint that the regions are closed meaning that there are no cut edges in the middle of a region. Correlation clustering is a natural clustering criteria because the number of clusters is not a hyper-parameter that must be hand tuned on an image by image basis. Instead it is a function of the potentials θ . We now illustrate this with two examples. Consider that θ is strictly positive; then the optimal partition cuts no edges meaning all super pixels are in the same region. Similarly if θ is strictly negative; then the optimal partition puts each super pixel in a separate region.

We now define closeness formally. We denote the set of cycles including any edge f as $S(f)$. For each cut edge f and each cycle c containing f , at least one edge on c in addition to f must be cut. These constraints above are called cycle inequalities. They are written as follows.

$$\sum_{e \in c-f} X_e \geq X_f \quad \forall [f; c \in S(f)] \quad (1)$$

Given the cycle inequalities we can now write the exact form of correlation clustering.

$$\min_{X \geq 0} \sum_e \theta_e X_e \quad (2)$$

$$\sum_{e \in c-f} X_e \geq X_f \quad \forall [f; c \in S(f)] \quad \text{AND} \quad X_e \leq 1 \quad \forall e$$

The form above is the MC LP relaxation. The integer programming version forces elements of X to be either 0 or 1. Solving Eq 2 is done by via cutting plane operations. The iteration consists of solving the LP and adding new constraints from the set of violated cycle inequalities. These constraints can be found using shortest path operations.

3 Derivation of the Dual of MC LP

We now take the dual of the MC LP (Eq 2). This will facilitate our comparison of the PlanarCC and MC LP relaxations. We first place the cycle inequalities in the objective using Lagrange multipliers ψ and δ .

$$\min_{X \geq 0} \max_{\psi \geq 0, \delta \geq 0} \sum_e \theta_e X_e + \sum_e \delta_e (X_e - 1) + \sum_{f, c \in S(f)} \psi_{cf} (X_f - \sum_{e \in c-f} X_e) \quad (3)$$

Now we reverse the order of the max-min in the LP. This does not alter the value of the combined objective as the primal and dual of any LP have the same value.

$$\max_{\psi \geq 0, \delta \geq 0} \min_{X \geq 0} \sum_e \theta_e X_e + \sum_e \delta_e (X_e - 1) + \sum_{f, c \in S(f)} \psi_{cf} (X_f - \sum_{e \in c-f} X_e) \quad (4)$$

We now rearrange the variables so that each primal variable is associated with its respective product. To ease notation we introduce $D(e)$. Here $D(e)$ refers to the set of pairs $[c, f]$ in which e is on the cycle c and $e \neq f$.

$$\max_{\psi \geq 0, \delta \geq 0} \min_{X \geq 0} \sum_e -\delta_e + (\theta_e + \delta_e + \sum_{c \in S(e)} \psi_{ce} + \sum_{[c, f] \in D(e)} -\psi_{cf}) X_e \quad (5)$$

We now convert the dual into another linear program called the Dual MC LP.

$$\begin{aligned} & \max_{\psi \geq 0, \delta \geq 0} \sum_e -\delta_e \\ & \theta_e + \delta_e + \sum_{c \in S(e)} \psi_{ce} - \sum_{[c, f] \in D(e)} \psi_{cf} \geq 0 \end{aligned} \quad (6)$$

Finally we alter the sign on δ_e and move it to the other side.

$$\begin{aligned} & \max_{\psi \geq 0, \delta \leq 0} \sum_e \delta_e \\ & \theta_e + \sum_{c \in S(e)} \psi_{ce} - \sum_{[c, f] \in D(e)} \psi_{cf} \geq \delta_e \end{aligned} \quad (7)$$

Consider that we have solved for dual MC LP. We now apply the following additional operation to δ and ψ . We now seek to alter the form of the objective without altering its value until we satisfy the following property which will be useful in our analysis.

$$\min(\theta_e, 0) \leq \delta_e \quad (8)$$

Select any such e violating this property. Now select any $[c, f] \in D(e)$ such that $\psi_{cf} > 0$. One must exist otherwise δ_e would be greater than or equal to $\min(\theta_e, 0)$. Let $\epsilon \leftarrow \min(-\delta_e, \psi_{cf})$. Now set $\delta_e \leftarrow \delta_e + \epsilon$ and set $\psi_{cf} \leftarrow \psi_{cf} - \epsilon$. We have now increased the component of the objective corresponding to δ_e and decreased the component of the objective corresponding to δ_f by no more than ϵ . In fact the objective must be unaffected as otherwise the LP would not be optimal. We continue this process until the property in Eq 8 is satisfied. This process can not continue forever because eventually all of the Lagrange multipliers ψ_{cf} would converge to zero forcing termination. Our final LP is as follows.

$$\begin{aligned} & \max_{\psi \geq 0, \delta \leq 0} \sum_e \delta_e \\ & \theta_e + \sum_{c \in S(e)} \psi_{ce} - \sum_{c, f \in D(e)} \psi_{cf} \geq \delta_e \quad \text{AND} \quad \delta_e \geq \min(0, \theta_e) \forall e \end{aligned} \quad (9)$$

We now interpret this as a dual decomposition over sub-problems. There is one sub-problem for each individual edge. The sub-problem corresponding to edge e is associated with potential δ_e . The MAP energy of the subproblem over edge e is simply δ_e . There is also one cycle sub-problem for each cycle c and edge on that cycle e . The potentials for that sub-problem equal ψ_{ce} for each edge $f \neq e$ on the cycle and $-\psi_{ce}$ for edge e . Notice that the MAP energy for each of the cycle sub-problems is zero.

4 Derivation of PlanarCC

The PlanarCC objective considers the matrix defining the set of all possible 2-colorable partitions Z . A 2-colorable partition is a partition in which each region can be given one of 2 colors such that it does not share a color with any of its neighbors. 2-colorable does not mean only two regions. For

example a checkerboard is a 2-colorable graph with 64 regions. Here each 2-colorable partition is associated with a single column of Z . Each row of Z corresponds to an edge e . We define Z to be a binary matrix where $Z_{e,r} = 1$ if and only if edge e is cut in partition r .

PlanarCC constructs a partition by taking a non-negative weighted sum of the columns of Z . This partition is defined using the non-negative column vector γ . Here γ has one index for each column of Z . The partition given by PlanarCC is $Z\gamma$. We also enforce that the benefit of cutting a negative edge e can not be less than θ_e . This is achieved using terms β and ϕ . We use β to indicate how much each edge is cut beyond the value 1. We use ϕ to denote the penalty for cutting an edge more than the value of 1. Here ϕ and β are indexed by e where $\phi_e = -\min(0, \theta_e)$. This means that if an edge e is over-cut ($(Z\gamma)_e > 1$) the penalty received is linear in the θ_e if the edge weight is negative and is zero otherwise.

Given this we write the objective of PlanarCC (primal problem).

$$\begin{aligned} \min_{\gamma \geq 0, \beta \geq 0} \quad & \theta Z\gamma + \phi\beta \\ & Z\gamma \leq 1 + \beta \end{aligned} \tag{10}$$

Solving Eq 10 is difficult because of the exponential size of Z . To circumvent this we take the dual problem, and solve via a cutting plane method. This produces a subset of Z such that no additional constraints are needed. We can then solve the primal LP by using this subset of Z as discussed later. We now derive the dual version of PlanarCC. We first convert the constraint into Lagrange multipliers λ .

$$\min_{\gamma \geq 0, \beta \geq 0} \max_{\lambda \geq 0} \theta Z\gamma + \phi\beta + \lambda(Z\gamma - 1 - \beta) \tag{11}$$

Now we reverse the order of the max-min in the LP.

$$\max_{\lambda \geq 0} \min_{\gamma \geq 0, \beta \geq 0} \theta Z\gamma + \phi\beta + \lambda(Z\gamma - 1 - \beta) \tag{12}$$

We now alter the form of the expression so as to group by primal variable.

$$\max_{\lambda \geq 0} \min_{\gamma \geq 0, \beta \geq 0} (\theta Z + \lambda Z)\gamma + (\phi - \lambda)\beta - \lambda 1 \tag{13}$$

We now convert the primal variables into constraints in the dual.

$$\begin{aligned} \max_{\lambda \geq 0} \quad & -\lambda 1 \\ & (\theta + \lambda)Z \geq 0 \\ & \phi \geq \lambda \end{aligned} \tag{14}$$

Now one should first notice that all non-negative entries of θ are associated with 0 valued Lagrange multipliers. Next one should notice that we are still troubled by the exponential size Z . We employ a cutting plane approach to solve Eq 14. In this manner we iteratively solve the LP then add the most violated constraint (which corresponds to a column of Z) to the LP. We treat finding the most violated constraint as MAP inference on binary planar ising model without a field Yarkony et al. (2012). This is commonly known as a binary planar Markov random field (MRF) with no unary (biases) potentials. In fact the solution to the binary planar ising model without a field is exactly the optimal 2-colorable partition given its potentials. MAP inference here can be done in $O(N^{\frac{3}{2}} \log N)$ time via a reduction to perfect matching Fisher (1966).

Once the LP has been solved the dual LP can be converted to a solution to the primal LP by solving the primal using only the Z identified during dual optimization. Alternatively one can produce an integer solution easily by taking the set of tight constraints over Z and then create a partition by taking the union edges cut in their respective partitions.

While the underlying problem is NP hard; for the data sets considered in recent work the LP is generally tight and when not is nearly tight. Furthermore the solutions produced are either exact or on rare occasions nearly exact Yarkony et al. (2012).

5 Relationship to PlanarCC

We now demonstrate that the PlanarCC LP and the MC LP have the same value. First we take the optimal MC LP dual solution and demonstrate that we can map this to a solution to the PlanarCC dual with the same value. This shows that dual PlanarCC LP is no less than the dual MC LP solution. Next take the optimal PlanarCC dual solution demonstrate that we can map it to a solution to the MC LP dual with the same objective value. This shows that dual PlanarCC LP no more than the dual MC LP solution. By certifying that the maximum value of dual PlanarCC is upper and lower bounded by the maximum value of the MC LP dual we show that the two must be equal.

5.1 Dual(PlanarCC) \geq Dual(MC)

Let δ correspond to the optimal solution to the lower bound in the MC LP relaxation. Let $\lambda \leftarrow -\delta$. Notice that if λ satisfies the constraints on the PlanarCC dual LP then the objective value of the dual PlanarCC LP is lower bounded by that of MC LP dual.

Observe that the inequality $0 \leq \lambda \leq \phi$ is satisfied. We now lower bound the amount the most violated constraint of the remaining constraints is violated by to zero.

$$\min_{X \in C_2} (\theta + \lambda)X \geq \min_{X \in C_2} \left(\sum_{c \in S(e)} -\psi_{ce} + \sum_{c, f \in D(e)} \psi_{cf} \right) X_e \quad (15)$$

Observe that if for any solution X , and $[c, f]$ if $-\psi_{cf}$ is contributed to the objective then ψ_{cf} must also be contributed to the objective at least once. This is because if edge f is ever cut then at least one additional edge on the cycle c must also be cut because $X \in C_2$. Thus the total sum in the lower bound in Eq 15 must be non-negative. Since cutting no edges is a valid value for X then the $\min_{X \in C_2} (\theta + \lambda)X = 0$.

Since $-\delta$ is a valid solution to the PlanarCC LP and has the objective same value as in MCLP then PlanarCC is no less than MCLP.

5.2 Dual(PlanarCC) \leq Dual(MC)

To prove that the value of the dual of PlanarCC is upper bounded by the dual of MC we rely on a tight dual formulation of MAP inference in a planar binary ising model without a field. Every binary planar ising model without a field has a dual LP with exactly the same value Barahona and Mahjoub. (1986) Yarkony et al. (2012). The form of this dual is as follows.

$$\min_{X \in C_2} (\theta + \lambda)X = \max_{\sum_c \omega_e^c = \lambda_e + \theta_e \forall e} \sum_c \min_{X^c \in C_2} \omega^c X^c \quad (16)$$

Here each vector ω^c is associated with one cycle of the graph. We use X^c to denote the minimizer over the set of 2-colorable partitions for $\omega^c X^c$. Since we know that $(\theta + \lambda)Z \geq 0$ and since every X^c is independently solved for all c then $\min_{X^c \in C_2} \omega^c X^c = 0$ then it must be that case that there is no more than one negative edge on any cycle ω^c . Moreover that negative edge must be have magnitude no greater than any other edge on the cycle. We write these constraints formally as follows.

$$(\omega_e^c < 0) \rightarrow (\omega_f^c + \omega_e^f \geq 0) \quad \forall c, f \in c, e \in c, e \neq f \quad (17)$$

Now consider the solution to the MCLP such that $\delta \leftarrow -\lambda$ and $\psi_{ce} = \max[0, -\omega_e^c]$. We consider that cycles c that include edge e and in which $\omega_e^c > 0$ to be in $D(e)$. We consider any c that includes e and in which $\omega_e^c < 0$ as in $S(e)$. Notice that $\psi_{cf} \leq \omega_e^c$ for all $[c, f] \in D(e)$.

We now establish that $[\delta, \psi]$ is an admissible solution to the LP in Eq 7 which has the equivalent to the final dual MC LP in Eq 9. We establish this by construction.

Claim: $\theta_e + \sum_{c \in S(e)} \psi_{ce} - \sum_{c, f \in D(e)} \psi_{cf} \geq \delta_e$

Proof:

We begin by writing $\theta + \lambda$ in terms of ω . We then swap out some terms in ω for equal terms in ψ .

$$\theta_e + \lambda_e = \theta_e - \delta_e = \sum_c \omega_e^c = \sum_{c \in S(e)} -\psi_{ce} + \sum_{c, f \in D(e)} \omega_f^c \quad (18)$$

Recall that $\psi_{cf} \leq \omega_e^c$ for all $[c, f] \in D(e)$. We now produce an inequality.

$$\theta_e - \delta_e \geq \sum_{c \in S(e)} -\psi_{ce} + \sum_{c, f \in D(e)} \psi_{cf} \quad (19)$$

We now rearrange the variables and recover our claim.

$$\theta_e + \sum_{c \in S(e)} \psi_{ce} - \sum_{c, f \in D(e)} \psi_{cf} \geq \delta_e \quad (20)$$

5.3 Final Statement

Since $\text{Dual}(\text{PlanarCC}) \leq \text{Dual}(\text{MC})$ and $\text{Dual}(\text{PlanarCC}) \geq \text{Dual}(\text{MC})$ then $\text{Dual}(\text{PlanarCC}) = \text{Dual}(\text{MC})$. Since the primal and dual forms of an LP have the same value then $\text{Primal}(\text{PlanarCC}) = \text{Primal}(\text{MC})$

6 Conclusion

Correlation clustering is a fascinating new direction at the intersection of graphical models and computer vision/image segmentation. PlanarCC is a powerful approach around which a family of algorithms have been formed. We provide an analysis of the PlanarCC LP relaxation and the MC LP relaxation and demonstrated that they correspond to the same LP.

References

- Nikhil Bansal, Avrim Blum, and Shuchi Chawla. Correlation clustering. In *Journal of Machine Learning*, pages 238–247, 2002.
- David R. Martin, Charless C. Fowlkes, and Jitendra Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(5):530–549, may 2004.
- Bjoern Andres, Joerg H. Kappes, Thorsten Beier, Ullrich Kothe, and Fred A. Hamprecht. Probabilistic image segmentation with closedness constraints. In *Proceedings of the Fifth International Conference on Computer Vision (ICCV-11)*, pages 2611–2618, 2011.
- Shai Bagon and Meirav Galun. Large scale correlation clustering 816 optimization. In *CoRR, abs/1112.2903*, 2011.
- Julian Yarkony, Alexander Ihler, and Charless Fowlkes. Fast planar correlation clustering for image segmentation. In *Proceedings of the 12th European Conference on Computer Vision (ECCV 2012)*, 2012.
- Yoram Bachrach, Pushmeet Kohli, Vladimir Kolmogorov, and Morteza Zadimoghaddam. Optimal coalition structures in graph games. *CoRR, abs/1108.5248*, 2011.
- Michael E. Fisher. On the dimer solution of planar ising models. *Journal of Mathematical Physics*, 7(10): 1776–1781, 1966.
- F Barahona and A Mahjoub. On cuts and matchings in planar graphs. *Mathematical Programming*, 60(1-3): 157–173, september 1986.