Supplemental material for the paper "Clustering via LP-based Stabilities"

Nikos Komodakis
University of Crete
komod@csd.uoc.gr

Nikos Paragios
Ecole Centrale de Paris
INRIA Saclay Ile-de-France
nikos.paragios@ecp.fr

Georgios Tziritas
University of Crete
tziritas@csd.uoc.gr

Abstract

This document contains technical proofs for all lemmas and theorems that are mentioned in the above referenced paper.

1 Appendix

Lemma 1. *Let* $\mathbf{h}$ *be an optimal dual solution to* DUAL*.*

1. *If* $\Delta_q(\mathbf{h}) > 0$ *then* $S(q) \geq \Delta_q(\mathbf{h})$*.*
2. *If* $\Delta_q(\mathbf{h}) < 0$ *then* $S(q) \leq \Delta_q(\mathbf{h})$*.*

Proof. For proving some of the theorems or lemmas of this paper, we will often need to vary the penalty d_{qq} that is associated with object q. For this reason, if that penalty takes the value z, we will hereafter denote the corresponding pair of primal and dual LP relaxations as PRIMAL(z) and DUAL(z) respectively (e.g., according to this notation it holds PRIMAL=PRIMAL(d_{qq}), DUAL=DUAL(d_{qq})). With a slight abuse of notation, we will hereafter denote the margin of a feasible solution to any problem DUAL(z) by $\Delta_q(\mathbf{h})$. Note, however, that $\Delta_q(\mathbf{h})$ depends on the value of the penalty z that is associated with q (of course, it will always be clear from context what that value is). We will also denote:

$$\Delta_q^+(\mathbf{h}) = \sum\nolimits_{p:h_{pq}=h_p} (\hat{h}_p - h_p) \ , \tag{1}$$

$$\Delta_q^-(\mathbf{h}) = \sum\nolimits_{p \neq q} (h_{pq} - \max(h_p, d_{pq})) + \Big(h_{qq} - h_q\Big) \ . \tag{2}$$

Obviously, it holds $\Delta_q^+(\mathbf{h}) \geq 0$, $\Delta_q^+(\mathbf{h}) \geq 0$ and also:

$$\Delta_q(\mathbf{h}) = \Delta_q^+(\mathbf{h}) - \Delta_q^-(\mathbf{h}) \ . \tag{3}$$

Furthermore, depending on whether $\Delta_q(\mathbf{h})$ is positive or negative, an optimal solution $\mathbf{h}$ must satisfy either $\Delta_q^+(\mathbf{h}) = 0$ or $\Delta_q^-(\mathbf{h}) = 0$ (otherwise one can easily prove that $\mathbf{h}$ can be modified such that its objective value increases by $\min(\Delta_q^+(\mathbf{h}), \Delta_q^-(\mathbf{h}))$).

Hence, if $\mathbf{h}$ is optimal to DUAL(d_{qq}) and satisfies $\Delta_q(\mathbf{h}) > 0$ it will hold $\Delta_q(\mathbf{h}) = \Delta_q^+(\mathbf{h})$. One can then easily update $\mathbf{h}$ into a feasible solution $\mathbf{h}'$ which satisfies $h'_{qq} = h'_q$, $\hat{h}'_q = h'_q + \Delta_q^+(\mathbf{h})$ and has the same objective value as solution $\mathbf{h}$, i.e., $\mathbf{h}'$ is also optimal to DUAL(d_{qq}). If we then set $\mathbf{h}'' = \mathbf{h}'$, $h''_{qq} = h'_{qq} + \Delta_q^+(\mathbf{h}) - \epsilon$ (where $0 < \epsilon < \Delta_q^+(\mathbf{h})$), the resulting $\mathbf{h}''$ would be an optimal solution to DUAL$(d_{qq} + \Delta_q^+(\mathbf{h}) - \epsilon)$ (since it is easy to show that it is feasible and satisfies all complementary slackness conditions). Furthermore, it holds $h''_{qq} = h''_q < \hat{h}''_q$, which, from complementary slackness, implies that there must exist optimal solution $\mathbf{x}$ to PRIMAL$(d_{qq} + \Delta_q^+(\mathbf{h}) - \epsilon)$ such that $x_{qq} > 0$.

It therefore holds $S(q) \geq \Delta_q^+(\mathbf{h}) - \epsilon = \Delta_q(\mathbf{h}) - \epsilon$ and, hence, $S(q) \geq \Delta_q(\mathbf{h})$ (since ϵ can be arbitrarily small).

Let us now consider the case where $\mathbf{h}$ is optimal to $\text{DUAL}(d_{qq})$ and satisfies $\Delta_q(\mathbf{h}) < 0$. It will then hold $\Delta_q(\mathbf{h}) = -\Delta_q^-(\mathbf{h})$. Hence, one can easily update $\mathbf{h}$ into a feasible solution $\mathbf{h}'$ which satisfies $h'_{qq} = h'_q + \Delta_q^-(\mathbf{h})$ and has the same objective value as solution $\mathbf{h}$, i.e., $\mathbf{h}'$ is also optimal to $\text{DUAL}(d_{qq})$. If we then set $\mathbf{h}'' = \mathbf{h}'$, $h''_{qq} = h'_{qq} - \Delta_q^-(\mathbf{h}) + \epsilon$ (where $0 < \epsilon < \Delta_q^-(\mathbf{h})$), the resulting $\mathbf{h}''$ would be an optimal solution to $\text{DUAL}(d_{qq} - \Delta_q^-(\mathbf{h}) + \epsilon)$ (since it is easy to show that it is feasible and satisfies all complementary slackness conditions). Furthermore, it holds $h''_{qq} > h''_q$, which, from complementary slackness, implies that there can be no optimal solution $\mathbf{x}$ to $\text{PRIMAL}(d_{qq} - \Delta_q^-(\mathbf{h}) + \epsilon)$ such that $x_{qq} > 0$. It therefore holds $S(q) \leq -\Delta_q^-(\mathbf{h}) + \epsilon = \Delta_q(\mathbf{h}) + \epsilon$ and, hence, $S(q) \leq \Delta_q(\mathbf{h})$ (since ϵ can be arbitrarily small).

□

Lemma 2.

1. *If there exists optimal solution* $\mathbf{x}$ *to* $\text{PRIMAL}(d_{qq})$ *such that* $x_{qq} > 0$*, then there exists optimal solution* $\mathbf{h}$ *to* $\text{DUAL}(d_{qq})$ *such that* $\Delta_q(\mathbf{h}) \geq 0$.
2. *Similarly, if there exists no optimal solution* $\mathbf{x}$ *to* $\text{PRIMAL}(d_{qq})$ *such that* $x_{qq} > 0$*, then there exists optimal solution* $\mathbf{h}$ *to* $\text{DUAL}(d_{qq})$ *such that* $\Delta_q(\mathbf{h}) \leq 0$.

Proof. Let $\mathbf{x}$ be an optimal solution to $\text{PRIMAL}(d_{qq})$ such that $x_{qq} > 0$. Let also $\mathbf{h}$ be an optimal solution to $\text{DUAL}(d_{qq})$ and let us assume that it satisfies $\Delta_q(\mathbf{h}) < 0$. Hence, as already explained in the proof of lemma 1, it will hold $\Delta_q(\mathbf{h}) = -\Delta_q^-(\mathbf{h})$. One can then easily update $\mathbf{h}$ into a feasible solution $\mathbf{h}'$ which satisfies $h'_{qq} = h'_q + \Delta_q^-(\mathbf{h})$ and has the same objective value as solution $\mathbf{h}$, i.e., $\mathbf{h}'$ is also optimal to $\text{DUAL}(d_{qq})$. However, due to conditions $h'_{qq} > h'_q$ and $x_{qq} > 0$, the pair of optimal solutions $(\mathbf{x}, \mathbf{h}')$ violates complementary slackness, which leads to a contradiction.

Let us now assume that no optimal solution $\mathbf{x}$ to $\text{PRIMAL}(d_{qq})$ exists such that $x_{qq} > 0$. Let us also assume that $\mathbf{h}$ is an optimal solution to $\text{DUAL}(d_{qq})$ which satisfies $\Delta_q(\mathbf{h}) > 0$. As already explained in the proof of lemma 1, it will hold $\Delta_q(\mathbf{h}) = \Delta_q^+(\mathbf{h})$. One can then easily update $\mathbf{h}$ into a feasible solution $\mathbf{h}'$ which satisfies $h'_{qq} = h'_q$, $\hat{h}'_q = h'_q + \Delta_q^+(\mathbf{h})$ and has the same objective value as solution $\mathbf{h}$, i.e., $\mathbf{h}'$ is also optimal to $\text{DUAL}(d_{qq})$. However, the condition $h'_{qq} = h'_q < \hat{h}'_q$ along with the fact that no optimal $\mathbf{x}$ exists such that $x_{qq} > 0$, imply that at least one complementary slackness condition will always be violated, which again leads to a contradiction. □

Theorem 3. *The following equalities hold true:*

$$S(q) \geq 0 \Rightarrow S(q) = \sup\{\Delta_q(\mathbf{h}), \mathbf{h} \textit{ optimal solution to } \text{DUAL}\} , \tag{4}$$

$$S(q) \leq 0 \Rightarrow S(q) = \inf\{\Delta_q(\mathbf{h}), \mathbf{h} \textit{ optimal solution to } \text{DUAL}\} . \tag{5}$$

Furthermore, it can be shown that:

$$S(q) = \textit{sign}(S(q)) \cdot \sup\{|\Delta_q(\mathbf{h})|, \mathbf{h} \textit{ optimal solution to } \text{DUAL}\} . \tag{6}$$

Proof. We denote:

$$S^+(q) = \sup\{\Delta_q(\mathbf{h}), \mathbf{h} \text{ optimal solution to DUAL}\} , \tag{7}$$

$$S^-(q) = \inf\{\Delta_q(\mathbf{h}), \mathbf{h} \text{ optimal solution to DUAL}\} . \tag{8}$$

Let us first consider the case where there exists optimal dual solution $\mathbf{h_0}$ to $\text{DUAL}(d_{qq})$ such that $\Delta_q(\mathbf{h_0}) > 0$. We will then show that $S(q) = S^+(q)$. Obviously, due to $\Delta_q(\mathbf{h_0}) > 0$, it will hold $S^+(q) > 0$ and, so, by definition of $S^+(q)$, there must exist arbitrarily small $\epsilon \geq 0$ and optimal solution $\mathbf{h}$ to $\text{DUAL}(d_{qq})$ such that $\Delta_q(\mathbf{h}) = S^+(q) - \epsilon > 0$. By lemma 1 above, it then follows that $S(q) \geq \Delta_q(\mathbf{h}) = S^+(q) - \epsilon$, which implies (due to ϵ being either arbitrarily small or zero):

$$S(q) \geq S^+(q) . \tag{9}$$

In this case it must, of course, hold $S(q) > 0$ as well.

Also, by definition of $S(q)$, there must exist arbitrarily small $\epsilon \geq 0$ and optimal solution $\mathbf{x}$ to $\textsc{Primal}(d_{qq} + S(q) - \epsilon)$ such that $x_{qq} > 0$. By lemma 2 above, this means that there must exist optimal $\mathbf{h}$ to $\textsc{Dual}(d_{qq} + S(q) - \epsilon)$ such that $\Delta_q(\mathbf{h}) \geq 0$. Since $S(q) - \epsilon > 0$ (due to that $S(q) > 0$ and ϵ is arbitrarily small), it is then easy to construct an optimal solution $\mathbf{h}'$ to $\textsc{Dual}(d_{qq})$ such that:

$$\Delta_q(\mathbf{h}') \geq S(q) - \epsilon \tag{10}$$

($\mathbf{h}'$ can be constructed from $\mathbf{h}$ by appropriately decreasing those pseudo-distances h_{pq} for which $h_{pq} = h_p$ while also ensuring that complementary slackness conditions hold true for $\mathbf{h}'$). From (10) it follows that $S^+(q) \geq S(q) - \epsilon$, which implies (due to that ϵ has to be either arbitrarily small or zero):

$$S^+(q) \geq S(q)\ . \tag{11}$$

From (9),(11), we conclude that $S(q) = S^+(q)$.

Let us now consider the case where there exists optimal $\mathbf{h_0}$ to $\textsc{Dual}(d_{qq})$ such that $\Delta_q(\mathbf{h_0}) < 0$. We will then show that $S(q) = S^-(q)$. Obviously, due to $\Delta_q(\mathbf{h_0}) < 0$, it will hold $S^-(q) < 0$ and, so, by definition of $S^-(q)$, there must exist arbitrarily small $\epsilon \geq 0$ and optimal solution $\mathbf{h}$ to $\textsc{Dual}(d_{qq})$ such that $\Delta_q(\mathbf{h}) = S^-(q) + \epsilon < 0$. By lemma 1 above, it then follows that $S(q) \leq \Delta_q(\mathbf{h}) = S^-(q) + \epsilon$, which implies (due to that ϵ has to be either arbitrarily small or zero):

$$S(q) \leq S^-(q)\ . \tag{12}$$

This, of course, also implies that $S(q)$ is negative in this case (i.e., $S(q) < 0$).

Also, by definition of $S(q)$, there must exist arbitrarily small $\epsilon \geq 0$ such that $\textsc{Primal}(d_{qq}+S(q)+\epsilon)$ has no optimal solution $\mathbf{x}$ with $x_{qq} > 0$. Hence, by lemma 2 above, there must exist optimal solution $\mathbf{h}$ to $\textsc{Dual}(d_{qq} + S(q) + \epsilon)$ such that $\Delta_q(\mathbf{h}) \leq 0$. Since $S(q) + \epsilon < 0$ (due to that $S(q) < 0$ and ϵ is arbitrarily small), it is then easy to construct optimal solution $\mathbf{h}'$ to $\textsc{Dual}(d_{qq})$ such that:

$$\Delta_q(\mathbf{h}') \leq S(q) + \epsilon \tag{13}$$

($\mathbf{h}'$ can be constructed from $\mathbf{h}$ by appropriately increasing those pseudo-distances h_{pq} for which $h_{pq} > h_p$ while also ensuring that complementary slackness conditions hold true for $\mathbf{h}'$). From (13) it follows that $S^-(q) \leq S(q) + \epsilon$, which implies (due to ϵ being either arbitrarily small or zero):

$$S^-(q) \leq S(q)\ . \tag{14}$$

From (12),(14), we conclude that $S(q) = S^-(q)$.

It remains to consider the case where $\Delta_q(\mathbf{h}) = 0$ for any optimal solution $\mathbf{h}$ to $\textsc{Dual}(d_{qq})$ (i.e., $S^+(q) = S^-(q) = 0$). In this case, using reductio ad absurdum, it is easy to show that for any $\epsilon > 0$ no optimal solution $\mathbf{x}$ to $\textsc{Primal}(d_{qq} + \epsilon)$ can satisfy $x_{qq} > 0$ as well as that for any $\epsilon > 0$ there always exists optimal solution $\mathbf{x}$ to $\textsc{Primal}(d_{qq} - \epsilon)$ such that $x_{qq} = 0$, thus proving that $S(q) = 0$.

□

Theorem 4. *If* $\max_{q \notin \mathcal{Q}} \Delta_q(\mathbf{h}) < 0$*, then the* DISTRIBUTE *operation maintains feasibility and, unless* $\mathcal{V} = \mathcal{Q} \cup \mathcal{L}_\mathcal{Q}$*, it also strictly increases the dual objective.*

Proof. Let $\mathbf{h}$, $\mathbf{h}'$ denote respectively the dual solution before and after the operation DISTRIBUTE. Due to $\max_{q \notin \mathcal{Q}} \Delta_q(\mathbf{h}) < 0$, feasibility condition $h'_{pq} \geq d_{pq}$ is trivial to check. Therefore, to prove feasibility of $\mathbf{h}'$, it suffices to verify that condition $\sum_p h'_{pq} = \sum_p d_{pq}$ holds true for all $q \notin \mathcal{Q}$. Indeed:

$$\begin{aligned}
\sum_p h'_{pq} &= \sum_{p \in \mathcal{Q}} h'_{pq} + \sum_{p \notin \mathcal{Q} \cup \{q\} : h_p < d_{pq}} h'_{pq} + \sum_{p \neq q, p \in \mathcal{L}_\mathcal{Q} : h_p \geq d_{pq}} h'_{pq} + \sum_{p \in \mathcal{V}_q : h_{pq} > h_p} h'_{pq} + \sum_{p \in \mathcal{V}_q : h_{pq} = h_p} h'_{pq} \\
&= \sum_{p \in \mathcal{Q}} h_{pq} + \sum_{p \notin \mathcal{Q} \cup \{q\} : h_p < d_{pq}} \big(h_{pq} - (h_{pq} - d_{pq})\big) + \sum_{p \neq q, p \in \mathcal{L}_\mathcal{Q} : h_p \geq d_{pq}} \big(h_{pq} - (h_{pq} - h_p)\big) + \\
&\quad \sum_{p \in \mathcal{V}_q : h_{pq} > h_p} \Big(h_{pq} - (h_{pq} - h_p) - \frac{\Delta_q(\mathbf{h})}{|\mathcal{V}_q|}\Big) + \sum_{p \in \mathcal{V}_q : h_{pq} = h_p} \Big(h_{pq} + (\hat{h}_p - h_p) - \frac{\Delta_q(\mathbf{h})}{|\mathcal{V}_q|}\Big) \\
&= \big(\sum_p h_{pq}\big) + \Delta_q(\mathbf{h}) - |\mathcal{V}_q| \cdot \frac{\Delta_q(\mathbf{h})}{|\mathcal{V}_q|} = \sum_p h_{pq} = \sum_p d_{pq}
\end{aligned}$$

Also, it is trivial to verify that the DISTRIBUTE operation does not decrease any minimum pseudo-distance, i.e., it holds $h'_p \geq h_p$. Furthermore, if there exists $p \notin \mathcal{Q} \cup \mathcal{L}_{\mathcal{Q}}$, then DISTRIBUTE will strictly increase the minimum pseudo-distance h_p (e.g., if $h_{pq} = h_p$ then DISTRIBUTE will raise h_{pq} by $-\frac{\Delta_q(\mathbf{h})}{|\mathcal{V}_q|} > 0$).

□

Theorem 5. *If* $\max_{q \notin \mathcal{Q}} \Delta_q(\mathbf{h}) > 0$*, then the* EXPAND *operation strictly decreases the primal cost* $E(\mathcal{Q})$.

Proof. Let $\bar{q} = \arg\max_{q \notin \mathcal{Q}} \Delta_q(\mathbf{h})$. By assumption, it holds $\Delta_{\bar{q}}(\mathbf{h}) > 0$. It is then easy to show that the primal cost related to all objects in $p \in \mathcal{V}_{\bar{q}}$ will decrease if we choose $\bar{q}$ as a new cluster center. In particular, the primal cost of making $\bar{q}$ a cluster center and assigning to it each $p \in \mathcal{V}_{\bar{q}} - \{\bar{q}\}$ is equal to $\sum_{p \in \mathcal{V}_{\bar{q}}} h_{p\bar{q}}$, whereas assigning each $p \in \mathcal{V}_{\bar{q}}$ to one of the current cluster centers in $\mathcal{Q}$ has primal cost strictly greater than $\sum_{p \in \mathcal{V}_{\bar{q}}} h_{p\bar{q}}$. As a result even by merely making $\bar{q}$ an active center and assigning to it each $p \in \mathcal{V}_{\bar{q}} - \{\bar{q}\}$ is guaranteed to decrease the primal cost.

□