Dual Decomposition for Joint Discrete-Continuous Optimization Supplementary Material

Christopher Zach
Microsoft Research Cambridge

March 6, 2013

Contents

1 The Dual of $E_{\text{DC-MRF}}$

We recall the DC-MRF primal energy with pairwise potentials.

$$E_{\text{DC-MRF}}(x,y) = \sum_{(s,t)\in\mathcal{E}} \sum_{i,j} (f_{st}^{ij})_{\oslash} \left(x_{st}^{ij}, y_{st\to s}^{ij}, y_{st\to t}^{ij} \right) \tag{1}$$
$$\text{s.t. } x_s^i = \sum_j x_{st}^{ij}, \qquad x_t^j = \sum_i x_{st}^{ij}, \qquad x_s \in \Delta^L,\, x_{st} \in \Delta^{L^2},$$
$$y_s^i = \sum_j y_{st\to s}^{ij}, \qquad y_t^j = \sum_i y_{st\to t}^{ij}$$

We absorbed any bounds on the arguments of f_{st}^{ij} into those functions. In order to derive (a particular) dual we need the following fact:

Fact 1. The conjugate of $\phi(x,y) \stackrel{\text{def}}{=} \sum_i (f_i)_{\oslash}(x_i,y_i) + \imath_\Delta(x)$ is given by

$$\phi^*(z,w) = \max_i \left\{ z_i + (f_i)^*(w_i) \right\}. \tag{2}$$

Proof. We need to calculate

$$\begin{aligned}
\phi^*(z,w) &= \max_{x\in\Delta,y} x^T z + y^T w - \sum_i x_i f_i(y_i/x_i) && [t_i = y_i/x_i] \\
&= \max_{x\in\Delta,t} \sum_i x_i \left(z_i + t_i w_i - f_i(t_i) \right) \\
&= \max_{x\in\Delta} \sum_i x_i \left(z_i + \max_{t_i} \{ t_i w_i - f_i(t_i) \} \right) \\
&= \max_{x\in\Delta} \sum_i x_i \big(z_i + f_i^*(w_i) \big) = \max_i \left\{ z_i + (f_i)^*(w_i) \right\}.
\end{aligned}$$

This concludes the proof. □

By introducing respective Lagrange multipliers $p^i_{st\to s}$, $p^i_{st\to t}$, $q^i_{st\to s}$, $q^i_{st\to t}$, we obtained the following dual

$$-E^*_{\text{DC-MRF}}(p,q) = \sum_{(s,t)\in\mathcal{E}} \max_{i,j} \left\{ -p^i_{st\to s} - p^j_{st\to t} + (f^{ij}_{st})^*(-q^i_{st\to s}, -q^j_{st\to t}) \right\} \tag{3}$$

$$+ \sum_s \max_i \left\{ \sum_{t\in out(s)} p^i_{st\to s} + \sum_{t\in in(s)} p^i_{ts\to s} \right\} + \sum_{s,i} \imath \left\{ \sum_{t\in out(s)} q^i_{st\to s} + \sum_{t\in in(s)} q^i_{ts\to s} = 0 \right\}.$$

The last two terms come from $\sum_s \imath_\Delta(x_s) + \sum_s 0^T y_s$ in the primal.

2 The Convex Relaxation Derived from $L_{\text{DD-I}}$

After introducing Lagrange multipliers $\lambda^i_{st\to s}$ and $\lambda^i_{st\to t}$ for the consistency constraints between x_s and x_{st}, we have the following Lagrangian:

$$\begin{aligned}
L_{\text{DD-I}}(x,z;\lambda,\mu) &= \sum_{s\sim t}\sum_{i,j} x^{ij}_{st}\left(f^{ij}_{st}(z_{st\to s}, z_{st\to t}) - \lambda^i_{st\to s} - \lambda^j_{st\to t}\right) + \sum_{s,i} x^i_s \left(\sum_{t\in out(s)} \lambda^i_{st\to s} + \sum_{t\in in(s)} \lambda^i_{ts\to s} \right) \\
&\quad + \sum_{s\sim t} \left(\mu_{st\to s}(z_s - z_{st\to s}) + \mu_{st\to t}(z_t - z_{st\to t}) \right) \\
&= \sum_{s\sim t} \left(\sum_{i,j} x^{ij}_{st}\left(f^{ij}_{st}(z_{st\to s}, z_{st\to t}) - \lambda^i_{st\to s} - \lambda^j_{st\to t}\right) - \mu^T_{st} z_{st} \right) \\
&\quad + \sum_{s,i} x^i_s \left(\sum_{t\in out(s)} \lambda^i_{st\to s} + \sum_{t\in in(s)} \lambda^i_{ts\to s} \right) + \sum_s z_s \left(\sum_{t\in out(s)} \mu_{st\to s} + \sum_{t\in in(s)} \mu_{ts\to s} \right).
\end{aligned}$$

Since we minimize over $x_s \in \Delta^L$ and $x_{st} \in \Delta^{L^2}$, we eliminate x and obtain

$$\begin{aligned}
L_{\text{DD-I}}(z;\lambda,\mu) &= \sum_{s\sim t} \left(\min_{i,j} \left\{ f^{ij}_{st}(z_{st\to s}, z_{st\to t}) - \lambda^i_{st\to s} - \lambda^j_{st\to t} \right\} - \mu^T_{st} z_{st} \right) + \sum_s \min_i \left\{ \sum_{t\in out(s)} \lambda^i_{st\to s} + \sum_{t\in in(s)} \lambda^i_{ts\to s} \right\} \\
&\quad + \sum_s z_s \left(\sum_{t\in out(s)} \mu_{st\to s} + \sum_{t\in in(s)} \mu_{ts\to s} \right).
\end{aligned}$$

Elimination of z by minimizing over $z_s, z_{st} \in \mathbb{R}$ yields the dual

$$\begin{aligned}
E^*_{\text{DD-I}}(\lambda,\mu) &= \min_{z_s, z_{st}} \sum_{s\sim t} \left(\min_{i,j} \left\{ f^{ij}_{st}(z_{st}) - \lambda^i_{st\to s} - \lambda^j_{st\to t} \right\} - \mu^T_{st} z_{st} \right) \\
&\quad + \sum_s \min_i \left\{ \sum_{t\in out(s)} \lambda^i_{st\to s} + \sum_{t\in in(s)} \lambda^i_{ts\to s} \right\} + \sum_s z_s \left(\sum_{t\in out(s)} \mu_{st\to s} + \sum_{t\in in(s)} \mu_{ts\to s} \right) \\
&= \sum_{s\sim t} \min_{z_{st}} \left\{ \min_{i,j} \left\{ f^{ij}_{st}(z_{st}) - \lambda^i_{st\to s} - \lambda^j_{st\to t} \right\} - \mu^T_{st} z_{st} \right\} \\
&\quad + \sum_s \min_i \left\{ \sum_{t\in out(s)} \lambda^i_{st\to s} + \sum_{t\in in(s)} \lambda^i_{ts\to s} \right\} + \sum_s \min_{z_s} \left\{ z_s \left(\sum_{t\in out(s)} \mu_{st\to s} + \sum_{t\in in(s)} \mu_{ts\to s} \right) \right\} \\
&= \sum_{s\sim t} - \max_{z_{st}} \left\{ \mu^T_{st} z_{st} - \min_{i,j} \left\{ f^{ij}_{st}(z_{st}) - \lambda^i_{st\to s} - \lambda^j_{st\to t} \right\} \right\}
\end{aligned}$$

$$+\sum_s \min_i \left\{ \sum_{t\in out(s)} \lambda^i_{st\to s} + \sum_{t\in in(s)} \lambda^i_{ts\to s} \right\} - \sum_s \imath \left\{ \sum_{t\in out(s)} \mu_{st\to s} + \sum_{t \in in(s)} \mu_{ts\to s} = 0 \right\}$$

$$= \sum_{s\sim t} -\max_{ij} \left\{ (f^{ij}_{st})^*(\mu_{st}) + \lambda^i_{st\to s} + \lambda^j_{st\to t} \right\}$$

$$+\sum_s \min_i \left\{ \sum_{t\in out(s)} \lambda^i_{st\to s} + \sum_{t\in in(s)} \lambda^i_{ts\to s} \right\} - \sum_s \imath \left\{ \sum_{t\in out(s)} \mu_{st\to s} + \sum_{t \in in(s)} \mu_{ts\to s} = 0 \right\},$$

where we used tha fact that $(\inf_i f_i)^* = \sup_i f_i^*$ (see [Hiriart-Urruty & Lemarechal, Thm 2.4.1]). In order to compute the primal we rewrite $L(\lambda, \mu)$ as

$$E^*_{\text{DD-I}}(\lambda,\mu) = \sum_{(s,t)\in\mathcal{E}} -\max_{ij} \left\{ (f^{ij}_{st})^*(\mu^{ij}_{st}) + \lambda^{ij}_{st} \right\} - \sum_s \max_i \lambda^i_s - \sum_s \imath\left\{\mu_s = 0\right\} + \sum_{(s,t),i} 0\cdot\lambda^i_{st\to s} + \sum_{(s,t)} 0^T \mu_{st}$$

subject to (we state the corresponding multipliers in the right column)

$$\begin{aligned}
\lambda^{ij}_{st} &= \lambda^i_{st\to s} + \lambda^j_{st\to t} && [x^{ij}_{st}]\\
\lambda^i_s &= -\sum_{t\in out(s)} \lambda^i_{st\to s} - \sum_{t\in in(s)} \lambda^i_{ts\to s} && [x^i_s]\\
\mu^{ij}_{st} &= \mu_{st} && [z^{ij}_{st}]\\
\mu_s &= -\sum_{t\in out(s)} \mu_{st\to s} - \sum_{t\in in(s)} \mu_{ts\to s}. && [z_s]
\end{aligned}$$

We explicitly added the zero terms on the additional unknowns to highlight that they correspond to constraints in the primal. $0\cdot\lambda^i_{st\to s}$ translates to the usual marginalization constraints, $x^i_s = \sum_j x^{ij}_{st}$ etc. $0\cdot\mu_{st\to s}$ e.g. translates to $z_s = \sum_{ij} z^{ij}_{st\to s}$, since $\mu_{st\to s}$ appears with $+1$ in constraints z_s and with -1 in constraints z^{ij}_{st} for all i, j. Hence, the corresponding primal reads as

$$\begin{aligned}
E_{\text{DC-DD-I}}(x,z) = \sum_{s,t}\sum_{i,j} (f^{ij}_{st})_\oslash (x^{ij}_{st}, z^{ij}_{st}) \qquad (4)\\
\text{s.t. } x^i_s = \sum_j x^{ij}_{st}, \qquad x^j_t = \sum_i x^{ij}_{st}, \qquad x_s \in \Delta^L,\ x_{st}\in\Delta^{L^2}\\
z_s = \sum_{i,j} z^{ij}_{st\to s}, \qquad z_t = \sum_{i,j} z^{ij}_{st\to t}.
\end{aligned}$$

3 The Convex Relaxation Derived from L_{DD}

Again, after introducing Lagrange multipliers $\lambda^i_{st\to s}$ and $\lambda^i_{st\to t}$ for the consistency constraints between x_s and x_{st}, we have the following Lagrangian:

$$\begin{aligned}
L_{\text{DD}}(x,z;\lambda,\mu) &= \sum_{s\sim t}\sum_{i,j} x^{ij}_{st}\left(f^{ij}_{st}(z^i_{st\to s}, z^j_{st\to t}) - \lambda^i_{st\to s} - \lambda^j_{st\to t}\right) + \sum_{s,i} x^i_s \left(\sum_{t\in out(s)} \lambda^i_{st\to s} + \sum_{t\in in(s)} \lambda^i_{ts\to s}\right)\\
&+ \sum_{s\sim t}\sum_i x^i_s \left(\mu^i_{st\to s}(z_s - z^i_{st\to s}) + \mu^i_{st\to t}(z_t - z^i_{st\to t})\right)\\
&= \sum_{s\sim t}\sum_{i,j} x^{ij}_{st}\left(f^{ij}_{st}(z^i_{st\to s}, z^j_{st\to t}) - \lambda^i_{st\to s} - \lambda^j_{st\to t} - \mu^i_{st\to s} z^i_{st\to s} - \mu^j_{st\to t} z^j_{st\to t}\right)
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{s,i} x_s^i \left(\sum_{t \in out(s)} \lambda^i_{st\to s} + \sum_{t \in in(s)} \lambda^i_{ts\to s} \right) + \sum_s z_s \sum_i x_s^i \left(\sum_{t \in out(s)} \mu^i_{st\to s} + \sum_{t \in in(s)} \mu^i_{ts\to s} \right) \\
&= \sum_{s\sim t} \sum_{i,j} x_{st}^{ij} \left(f_{st}^{ij}(z^i_{st\to s}, z^j_{st\to t}) - \lambda^i_{st\to s} - \lambda^j_{st\to t} - \mu^i_{st\to s} z^i_{st\to s} - \mu^j_{st\to t} z^j_{st\to t} \right) \\
&+ \sum_{s,i} x_s^i \left(\sum_{t \in out(s)} \lambda^i_{st\to s} + \sum_{t \in in(s)} \lambda^i_{ts\to s} + z_s \sum_{t \in out(s)} \mu^i_{st\to s} + z_s \sum_{t \in in(s)} \mu^i_{ts\to s} \right).
\end{aligned}
$$

In the second line we expanded the marginalization constraints, e.g. $x_s^i = \sum_j x_{st}^{ij}$, to move the terms into the first sum. In order to obtain the dual we minimize over x_s and x_{st} subject to the simplex constraints, and the dual energy is computed by minimizing over z,

$$
\begin{aligned}
E^*_{\text{DD}}(\lambda,\mu) &= \sum_{s\sim t} \min_{\{z^i_{st\to s}, z^j_{st\to s}\}} \min_{i,j} \left\{ f_{st}^{ij}(z^i_{st\to s}, z^j_{st\to t}) - \lambda^i_{st\to s} - \lambda^j_{st\to t} - \mu^i_{st\to s} z^i_{st\to s} - \mu^j_{st\to t} z^j_{st\to t} \right\} \\
&+ \sum_s \min_{z_s} \min_i \left\{ \sum_{t \in out(s)} \lambda^i_{st\to s} + \sum_{t \in in(s)} \lambda^i_{ts\to s} + z_s \sum_{t \in out(s)} \mu^i_{st\to s} + z_s \sum_{t \in in(s)} \mu^i_{ts\to s} \right\} \\
&= \sum_{s\sim t} \min_{i,j} \min_{z^i_{st\to s}, z^j_{st\to s}} \left\{ f_{st}^{ij}(z^i_{st\to s}, z^j_{st\to t}) - \lambda^i_{st\to s} - \lambda^j_{st\to t} - \mu^i_{st\to s} z^i_{st\to s} - \mu^j_{st\to t} z^j_{st\to t} \right\} \\
&+ \sum_s \min_i \min_{z_s} \left\{ \sum_{t \in out(s)} \lambda^i_{st\to s} + \sum_{t \in in(s)} \lambda^i_{ts\to s} + z_s \sum_{t \in out(s)} \mu^i_{st\to s} + z_s \sum_{t \in in(s)} \mu^i_{ts\to s} \right\} \\
&= \sum_{s\sim t} \min_{i,j} - \max_{z^i_{st\to s}, z^j_{st\to s}} \left\{ \mu^i_{st\to s} z^i_{st\to s} + \mu^j_{st\to t} z^j_{st\to t} - f_{st}^{ij}(z^i_{st\to s}, z^j_{st\to t}) + \lambda^i_{st\to s} + \lambda^j_{st\to t} \right\} \\
&+ \sum_s \min_i \left\{ \sum_{t \in out(s)} \lambda^i_{st\to s} + \sum_{t \in in(s)} \lambda^i_{ts\to s} - \imath \left\{ \sum_{t \in out(s)} \mu^i_{st\to s} + \sum_{t \in in(s)} \mu^i_{ts\to s} = 0 \right\} \right\} \\
&= \sum_{s\sim t} \min_{i,j} - \left\{ (f_{st}^{ij})^*(\mu^i_{st\to s}, \mu^j_{st\to t}) + \lambda^i_{st\to s} + \lambda^j_{st\to t} \right\} \\
&+ \sum_s \min_i \left\{ \sum_{t \in out(s)} \lambda^i_{st\to s} + \sum_{t \in in(s)} \lambda^i_{ts\to s} \right\} - \sum_{s,i} \imath \left\{ \sum_{t \in out(s)} \mu^i_{st\to s} + \sum_{t \in in(s)} \mu^i_{ts\to s} = 0 \right\} \\
&= - \sum_{s\sim t} \max_{i,j} \left\{ (f_{st}^{ij})^*(\mu^i_{st\to s}, \mu^j_{st\to t}) + \lambda^i_{st\to s} + \lambda^j_{st\to t} \right\} \\
&- \sum_s \max_i \left\{ - \sum_{t \in out(s)} \lambda^i_{st\to s} - \sum_{t \in in(s)} \lambda^i_{ts\to s} \right\} - \sum_{s,i} \imath \left\{ \sum_{t \in out(s)} \mu^i_{st\to s} + \sum_{t \in in(s)} \mu^i_{ts\to s} = 0 \right\} \\
&= E^*_{\text{DC-MRF}}(-\lambda, -\mu).
\end{aligned}
$$