Combinatorial Optimization in Computer Vision

Chapter 7: Multilabel Problems with Arbitrary Label Space: The Fast-PD Algorithm

WS 2011/12

Ulrich Schlickewei
Computer Vision and Pattern Recognition Group
Technische Universität München

Plan for Today

Introduction

Primal-Dual Principle

Overview of Fast-PD Algorithm

Details of Step 1, Step 2 and Step 3

Summary

MRFs with Arbitrary Label Set

Consider a Markov Random Field (MRF) $X = (X_v)_{v \in V}$ on a graph $G = (V, E)$ which takes values in an arbitrary label set $\mathcal{L}$.

Assume that the Gibbs potential of X is given by

$$E(x) = \sum_{v \in V} \phi_v(x(v)) + \sum_{v \sim w} a_{vw} \psi(x(v), x(w)).$$

Here, $\psi : \mathcal{L} \times \mathcal{L} \to \mathbb{R}$ is a semi-metric, a quasi-metric or a metric and $a_{vw} \geq 0$ are constants.

Again, we want to find the labelling with the least energy.

Fast PD versus Expansion Algorithm

- Today we are studying the Fast PD Algorithm by Komodakis, Tziritas and Paragios.

- This algorithm solves similar problems as the Expansion Algorithm with similar quality guarantees but with some advantages:
  - it is more efficient,
  - it is more general,
  - it is conceptually more elegant.

- We start by discussing the primal-dual principle in general and then we pass to the outline of the Fast-PD algorithm which essentially comprises three steps.

Plan for Today

Introduction

Primal-Dual Principle

Overview of Fast-PD Algorithm

Details of Step 1, Step 2 and Step 3

Summary

The Primal-Dual Principle I

Consider an Integer Linear Program (ILP)

$$\begin{aligned} \min \quad & c^\top x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \\ & x \in \mathbb{Z}^N. \end{aligned}$$

For simplicity we assume that $c \geq 0$. The dual LP corresponding to the LP relaxation is

$$\begin{aligned} \max \quad & b^\top y \\ \text{subject to} \quad & A^\top y \leq b. \end{aligned}$$

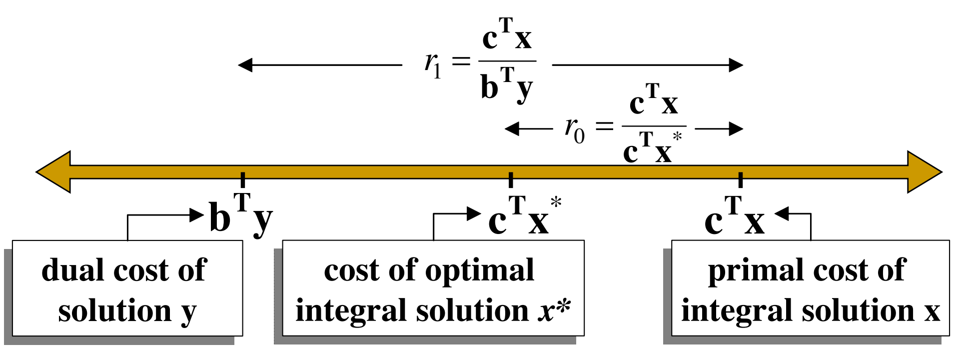
The Primal-Dual Principle II

Theorem: Let (x, y) be a pair of integral-primal and dual feasible solutions satisfying

$$c^\top x \leq f b^\top y$$

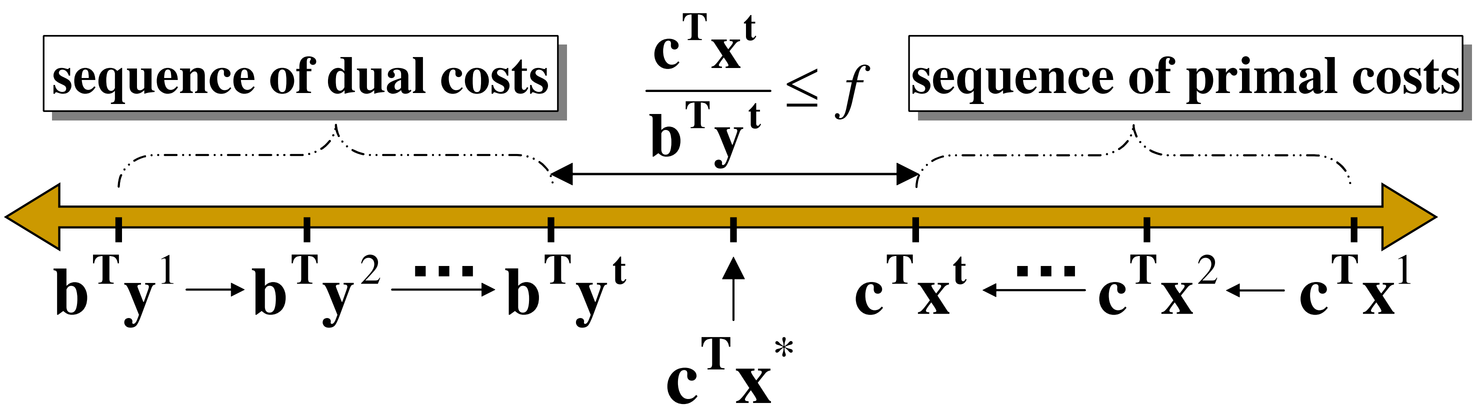
for a constant $f \geq 1$. Then x is an f - approximation of the global minimum x^* , that is,

$$c^\top x \leq f c^\top x^*.$$



The Primal-Dual Principle III

Typically, primal-dual algorithms work by constructing sequences $(x^k, y^k)_{k=1,\dots,K}$ of primal and dual solutions such that eventually the condition $c^\top x^K \leq f b^\top y^K$ is satisfied.



The Primal-Dual Principle IV

One possibility for constructing such a sequence $(x^k, y^k)_{k=1,\dots,K}$ which eventually fulfills the condition $c^\top x^K \leq f b^\top y^K$, relies on the complementary slackness principle. Let $\mathbf{f} = (f_j)_{j=1}^N \geq 1$ be a vector and let $f = \max_j (f_j)$.

Theorem (Relaxed Complementary Slackness Conditions): If a pair (x, y) of integral-primal and dual feasible solutions satisfies

$$\forall j: \left(x_j > 0\right) \implies \left(\sum_{i=1}^{M} a_{ij} y_i \geq \frac{c_j}{f_j}\right),$$

then (x, y) satisfies $c^\top x \leq f b^\top y$ and x is an f - approximation of the global minimum x^*.

The Primal-Dual Principle V

Proof: Using the fact that x is feasible, we have $Ax = b,\ x \geq 0.$
In combination with the condition

$$\forall j:\ \Big(x_j > 0\Big) \implies \Big(\sum_{i=1}^{M} a_{ij} y_i \geq \frac{c_j}{f_j}\Big).$$

this yields

$$\sum_j \frac{c_j}{f_j} x_j \leq x^\top (A^\top y) = (Ax)^\top y = b^\top y.$$

Since $c \geq 0$ by hypothesis, this implies that

$$c^\top x \leq \max_j (f_j) b^\top y.$$

Plan for Today

Introduction

Primal-Dual Principle

Overview of Fast-PD Algorithm

Details of Step 1, Step 2 and Step 3

Summary

Overview of the Fast-PD Algorithm: Step 1

Step 1: Reformulate the labelling problem as an Integer Linear Program (ILP)

$$\begin{aligned} \min \quad & c^\top x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \\ & x \in \mathbb{Z}^N. \end{aligned}$$

Determine the dual LP to the LP relaxation

$$\begin{aligned} \max \quad & b^\top y \\ \text{subject to} \quad & A^\top y \leq b. \end{aligned}$$

Overview of the Fast-PD Algorithm: Step 2

Step 2: Determine Relaxed Complementary Slackness conditions.

For that sake, fix a vector $\mathbf{f} = (f_j)_{j=1}^N \geq 1$ in $\mathbb{R}^N$ and set $f = \max_j (f_j)$. Then, if we are able to construct a pair (x, y) of integral-primal and dual feasible solutions which satisfy

$$\forall j : \left(x_j > 0\right) \implies \left(\sum_{i=1}^{M} a_{ij} y_i \geq \frac{c_j}{f_j}\right).$$

we can infer that x is an f - approximation to the optimal solution x^*.

Overview of the Fast-PD Algorithm: Step 3

Step 3: Keep generating pairs of feasible integral-primal and of dual solutions

$$(x^k, y^k)_{k=1,\ldots,K}$$

until both, x^K and y^K are feasible and satisfy the relaxed complementary slackness conditions

$$\left(x_j^K > 0\right) \implies \left(\sum_{i=1}^{M} a_{ij} y_i^K \geq \frac{c_j}{f_j}\right) \quad \forall j.$$

Then use x^K as an f - approximation to the unknown global solution.

Plan for Today

Introduction

Primal-Dual Principle

Overview of Fast-PD Algorithm

Details of Step 1, Step 2 and Step 3

Summary

ILP Formulation of Multilabelling Problem

The multilabel problem

$$x^* = \text{argmin}_x \sum_{v \in V} \phi_v(x(v)) + \sum_{v \sim w} a_{vw} \psi(x(v), x(w)).$$

can be formulated equivalently as the ILP

$$\begin{aligned}
\min_{x_{v,\alpha}, x_{\{(v,\alpha),(w,\beta)\}}} \quad & \sum_{v \in V, \alpha \in \mathcal{L}} c_{v,\alpha} x_v + \sum_{v \sim w, \alpha, \beta \in \mathcal{L}} a_{vw} \psi_{\alpha\beta} x_{\{(v,\alpha),(w,\beta)\}}, \\
\text{subject to} \quad & \sum_{\alpha \in \mathcal{L}} x_{v,\alpha} = 1 \;\; \forall \alpha \in \mathcal{L}, v \in V \\
& \sum_{\beta \in \mathcal{L}} x_{\{(v,\alpha),(w,\beta)\}} = x_{v,\alpha} \;\; \forall v \sim w, \alpha \in \mathcal{L} \\
& x_{v,\alpha}, x_{\{(v,\alpha),(w,\beta)\}} \in \{0, 1\} \;\; \forall v \sim w, \alpha, \beta \in \mathcal{L}.
\end{aligned}$$

LP Relaxation

In the LP relaxation of

$$\min_{x_{v,\alpha}, x_{\{(v,\alpha),(w,\beta)\}}} \quad \sum_{v \in V, \alpha \in \mathcal{L}} c_{v,\alpha} x_v + \sum_{v \sim w, \alpha, \beta \in \mathcal{L}} a_{vw} \psi_{\alpha\beta} x_{\{(v,\alpha),(w,\beta)\}},$$

$$\text{subject to} \quad \sum_{\alpha \in \mathcal{L}} x_{v,\alpha} = 1 \ \ \forall \alpha \in \mathcal{L}, v \in V$$

$$\sum_{\beta \in \mathcal{L}} x_{\{(v,\alpha),(w,\beta)\}} = x_{v,\alpha} \ \ \forall v \sim w, \alpha \in \mathcal{L}$$

$$x_{v,\alpha}, x_{\{(v,\alpha),(w,\beta)\}} \in \{0,1\} \ \ \forall v \sim w, \alpha, \beta \in \mathcal{L}.$$

the binary constraints $x_{v,\alpha}, x_{\{(v,\alpha),(w,\beta)\}} \in \{0,1\}$ are replaced by constraints $x_{v,\alpha}, x_{\{(v,\alpha),(w,\beta)\}} \geq 0$.

The Dual LP

The corresponding dual LP then reads as

$$\max_{y_v, y_{\{v,w\},(v,\alpha)}} \quad \sum_{v \in V} y_v$$

$$\text{subject to} \quad y_v - \sum_{v \sim w, \alpha \in \mathcal{L}} v_{\{v,w\},(v,\alpha)} \leq c_{v,\alpha} \quad \forall v \in V, \alpha \in \mathcal{L}$$

$$y_{\{v,w\},(v,\alpha)} + y_{\{v,w\},(w,\beta)} \leq a_{vw}\psi_{\alpha\beta} \quad \forall v \sim w, \alpha, \beta \in \mathcal{L}.$$

Plan for Today

Introduction

Primal-Dual Principle

Overview of Fast-PD Algorithm

Details of Step 1, Step 2 and Step 3

Summary

Parameters for Relaxed Complementary Slackness in Fast-PD

The Fast-PD algorithm as proposed by Komodakis, Tziritas and Paragios uses the (family of) relaxation vector(s)

$$\mathbf{f} = (\underbrace{\mu C, \ldots, \mu C}_{\text{first } |V \times \mathcal{L}| \text{ entries}}, \underbrace{C, \ldots, C}_{\text{last } 2|E \times \mathcal{L}| \text{ entries}}),$$

where $C = 2\frac{\max_{\alpha \neq \beta} \psi(\alpha, \beta)}{\min_{\alpha \neq \beta} \psi(\alpha, \beta)}$ and $\mu \in [\frac{1}{C}, 1]$ is a parameter.

Note that $\max(f_j) = C$ so that the Fast-PD algorithm will compute solutions with exactly the same quality guarantee as the expansion algorithm.

Relaxed Complementary Slackness Conditions

The relaxed complementary slackness conditions then read as

$$\Big(x_{v,\alpha} > 0\Big) \implies \Big(y_v - \sum_{v \sim w} y_{\{v,w\},(v,\alpha)} \geq \frac{c_{v,\alpha}}{\mu C}\Big) \qquad (*)$$

and

$$\Big(x_{\{(v,\alpha),(w,\beta)\}} > 0\Big) \implies \Big(y_{\{v,w\},(v,\alpha)} + y_{\{v,w\},(w,\beta)} \geq \frac{a_{vw}\psi_{\alpha\beta}}{C}\Big). \qquad (**)$$

Plan for Today

Introduction

Primal-Dual Principle

Overview of Fast-PD Algorithm

Details of Step 1, Step 2 and Step 3

Summary

Iterations of Fast-PD

The Fast-PD algorithm produces in each iteration a pair $(x^k, y^k)_{k=1,\dots,K}$ of feasible integral-primal and dual solutions such that (x^K, y^K) satisfy the two relaxed complementary slackness conditions

$$\left(x^K_{v,\alpha} > 0\right) \implies \left(y^K_v - \sum_{v \sim w} y^K_{\{v,w\},(v,\alpha)} \geq \frac{c_{v,\alpha}}{\mu C}\right) \qquad (*)$$

and

$$\left(x^K_{\{(v,\alpha),(w,\beta)\}} > 0\right) \implies \left(y^K_{\{v,w\},(v,\alpha)} + y^K_{\{v,w\},(w,\beta)} \geq \frac{a_{vw}\psi_{\alpha\beta}}{C}\right). \qquad (**)$$

Pseudo-Code

The outline of Fast-PD in Pseudo-Code is as follows:

1. $[x, y] \leftarrow \texttt{Init\_Duals\_Primals}()$
2. $x_{\texttt{old}} \leftarrow x$
3. for each $\alpha \in \mathcal{L}$

   $y \leftarrow \texttt{PreEdit\_Duals}(\alpha, x, y)$

   $[x', y'] \leftarrow \texttt{Update\_Duals\_Priamals}(\alpha, x, y)$

   $y' \leftarrow \texttt{PostEdit\_Duals}(\alpha, x', y')$

   $x' \leftarrow x;\ y' \leftarrow y$
4. if $(x \neq x_{\texttt{old}})$

   $x_{\texttt{old}} \leftarrow x$

   goto 3.

Comments

- In each iteration (Step 3.), the algorithm generates feasible integral-primal solutions and feasible dual solutions which satisfy the second slackness condition

$$\left(x_{\{(v,\alpha),(w,\beta)\}} > 0\right) \implies \left(y_{\{v,w\},(v,\alpha)} + y_{\{v,w\},(w,\beta)} \geq \frac{a_{vw}\psi_{\alpha\beta}}{C}\right). \qquad (**)$$

Each iteration leads to pairs (x, y) which are closer to satisfy the first slackness condition

$$\left(x_{v,\alpha} > 0\right) \implies \left(y_v - \sum_{v \sim w} y_{\{v,w\},(v,\alpha)} \geq \frac{c_{v,\alpha}}{\mu C}\right) \qquad (*)$$

Comments II

- The core step in the algorithm is hidden in the function $\texttt{Update\_Duals\_Primals}(\alpha, x, y)$. In this method, a max-flow through a graph is constructed which is similar to the graph used for computing the optimal α - expansion. The main difference is that the graph used for evaluating $\texttt{Update\_Duals\_Primals}(\alpha, x, y)$ admits provably less augmenting paths, therefore leading to an improved performance.
- If the parameter μ is set to 1, it can be shown that during each α-iteration (one iteration in the loop in step 3.) the optimal α - expansion is computed. Thus, in this case the algorithm produces the same result as the expansion algprotihm.

Plan for Today

Introduction

Primal-Dual Principle

Overview of Fast-PD Algorithm

Details of Step 1, Step 2 and Step 3

Summary

Summary

- The Primal-Dual principle is a general principle which is widely used in the literature for designing algorithms which approximate solutions to NP-hard problems.

- For the multilabel problem with pairwise interaction potential, the Fast-PD algorithm computes solutions with the same quality guarantee as the expansion algorithm.

- Each iteration of Fast-PD relies on the computation of a max-flow, which can be determined more efficiently than the computation of the optimal expansion.