# On Relation Between Constraint Propagation and Block-Coordinate Descent in Linear Programs

Tomáš Dlask, Tomáš Werner

Czech Technical University in Prague, Faculty of Electrical Engineering

{*dlaskto2,werner*}*@fel.cvut.cz*

24 January 2022, Heidelberg

# Block-Coordinate Descent (BCD)

- Iterative method, optimizes a multivariate function by solving block-variable subproblems while keeping the other variables constant
- Fixed points may not be global optima (even for convex problems)
- Convex message-passing algorithms
  - BCD applied to dual LP relaxation of Weighted CSP (a.k.a. MAP inference in graphical models/discrete energy minimization)
  - Fixed points characterized by local consistency conditions (arc consistency/node-edge agreement/...)

# Primal-Dual Approach With Constraint Propagation[2]

- Upper-bounding LPs
- Iteratively improves a dual solution by detecting infeasibility of complementary slackness conditions by constraint propagation
- Generalization of Virtual Arc Consistency algorithm[1]
- Fixed points also characterized by a local consistency condition

[1] Cooper et al.: Soft arc consistency revisited

[2] Dlask and Werner: Bounding Linear Programs by Constraint Propagation: Application to Max-SAT

# Overview of Results

- Block-coordinate descent
- Primal-dual approach

$\Big\}$ Identical fixed points

- Constraint propagation rule performed by BCD in any LP.
- BCD optimal $\iff$ propagation is refutation-complete.
- Characterization of LPs optimally solvable by BCD by tools of constraint programming.
- Characterization of types of local minima in BCD using local consistency conditions

# Optimization Problem

Primal-dual pair of linear programs:

$$\max c^T x \qquad \min b^T y \tag{1a}$$

$$Ax = b \qquad y \in \mathbb{R}^m \tag{1b}$$

$$x \geq 0 \qquad A^T y \geq c \tag{1c}$$

with $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

Suppose:

- dual-feasible solution $y$ is provided
- collection of subsets $\mathcal{B} \subseteq 2^{[m]}$ is given ($[m] = \{1, ..., m\}$ for brevity)

# Block-Coordinate Descent and Relative Interior Rule[3]

Recall the dual LP: $\min\{b^T y \mid A^T y \geq c, y \in \mathbb{R}^m\}$

Block-coordinate descent (input: feasible solution $y$):

1. Choose $B \in \mathcal{B}$ (e.g., cyclic choice)
2. Update $y$ such that $$y_B \in \operatorname*{argmin}_{y'_B \in \mathbb{R}^B}\{b_B^T y'_B \mid A^T(y'_B, y_{-B}) \geq c)\} \qquad (2)$$
3. Unless termination condition is satisfied, go to 1.

Relative interior rule: $$y_B \in \operatorname{ri} \operatorname*{argmin}_{y'_B \in \mathbb{R}^B}\{b_B^T y'_B \mid A^T(y'_B, y_{-B}) \geq c)\} \qquad (3)$$

**Definition:** A point $y$ feasible to the dual in (1) is

- a local minimum (LM) w.r.t. $\mathcal{B}$ if (2) holds $\forall B \in \mathcal{B}$,
- an interior local minimum (ILM) w.r.t. $\mathcal{B}$ if (3) holds $\forall B \in \mathcal{B}$,
- a pre-interior local minimum (pre-ILM) w.r.t. $\mathcal{B}$ if ... (see [3])

[3] Werner et al.: Relative interior rule in block-coordinate descent

# Primal-Dual Approach with Constraint Propagation[2]

$$\max c^T x \qquad\qquad \min b^T y$$
$$Ax = b \qquad\qquad y \in \mathbb{R}^m$$
$$x \geq 0 \qquad\qquad A^T y \geq c$$

- Define $J(y) = \{j \in [n] = \{1, ..., n\} \mid A_j^T y = c_j\}$ ($A_j$ is $j$-th column of $A$)
- Complementary slackness: dual-feasible $y$ is optimal if and only if

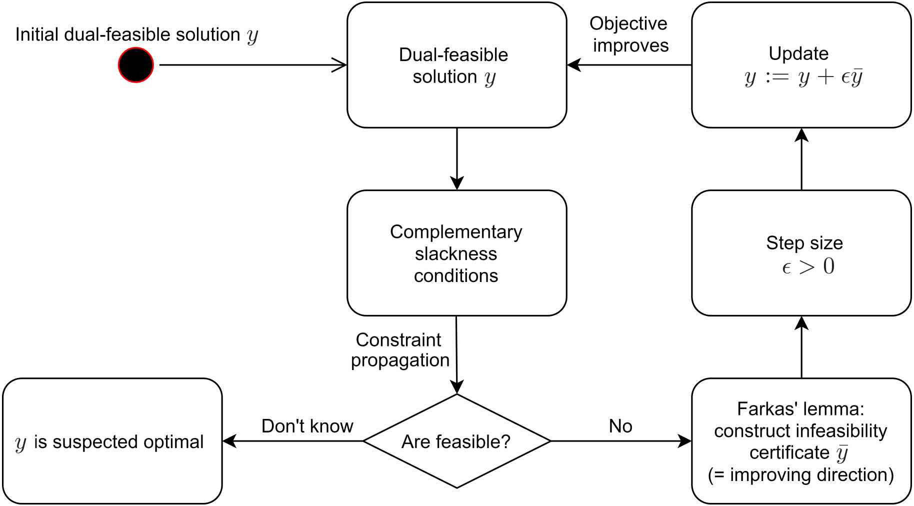
$$Ax = b \tag{5a}$$
$$x_j \geq 0 \qquad \forall j \in J(y) \tag{5b}$$
$$x_j = 0 \qquad \forall j \in [n] - J(y) \tag{5c}$$

is feasible.
- By Farkas' lemma, (5) is infeasible if and only if $\exists \bar{y} \in \mathbb{R}^m$:

$$b^T \bar{y} < 0 \tag{6a}$$
$$A_j^T \bar{y} \geq 0 \qquad \forall j \in J(y) \tag{6b}$$

[2] Dlask and Werner: Bounding Linear Programs by Constraint Propagation: Application to Max-SAT

# Primal-Dual Approach with Constraint Propagation[2]




[2] Dlask and Werner: Bounding Linear Programs by Constraint Propagation: Application to Max-SAT

# Considered Constraint Propagation Rule

**Rule:** *Choose a subset of (in)equalities, infer which hold with equality.*

Apply to the previously shown system:

$$Ax = b \tag{7a}$$
$$x_j \geq 0 \qquad \forall j \in J \tag{7b}$$
$$x_j = 0 \qquad \forall j \in [n] - J \tag{7c}$$

1. Initialize $J := J(y)$.
2. Choose a subset $B \in \mathcal{B}$, find all $j \in J$ such that
   $$A^i x = b_i \qquad \forall i \in B \subseteq [m] \tag{8a}$$
   $$x_j \geq 0 \qquad \forall j \in J \tag{8b}$$
   $$x_j = 0 \qquad \forall j \in [n] - J \tag{8c}$$
   implies $x_j = 0$.
   ($A^i$ is the $i$-th row of $A$)
3. Remove such indices from $J$.
4. Go to 2 until $J$ does not change or (8) is infeasible for some $B \in \mathcal{B}$.

# Constraint Propagation

$$A^i x = b_i \qquad \forall i \in B \tag{8a}$$
$$x_j \geq 0 \qquad \forall j \in J \tag{8b}$$
$$x_j = 0 \qquad \forall j \in [n] - J \tag{8c}$$

**Definition** (*propagator* $P_B \colon 2^{[n]} \to 2^{[n]} \cup \{\bot\}$): Let $B \subseteq [m]$, $J \subseteq [n]$.

$$P_B(J) = \begin{cases} \bot & \text{if (8) is infeasible} \\ J - \{j \in J \mid \text{(8) implies } x_j = 0\} & \text{otherwise} \end{cases}$$

**Definition:** For $B \subseteq [m]$, a set $J \subseteq [n]$ is *$B$-consistent* if (8) is feasible and for every $j \in J$ system (8) does not imply $x_j = 0$.

**Proposition:** If $J, J'$ are $B$-consistent, $J \cup J'$ is $B$-consistent.
$\to$ join-semilattice structure
$\to$ set $\{J \subseteq [n] \mid J \text{ is } B\text{-consistent}\} \cup \{\bot\}$ is a complete lattice
$\to$ $P_B$ is the (dual) closure operator associated with the complete lattice

# Constraint Propagation

*Propagation algorithm* is defined by $\mathcal{B} \subseteq 2^{[m]}$ and is given $J \subseteq [n]$:

1. Find $B \in \mathcal{B}$ with $P_B(J) \neq J$, update $J := P_B(J)$.
2. Repeat 1 until no such $B \in \mathcal{B}$ exists or $J = \bot$
3. Return $J$

**Definition:** For $\mathcal{B} \subseteq 2^{[m]}$, $J$ is $\mathcal{B}$-*consistent* if it is $B$-consistent $\forall B \in \mathcal{B}$.

Properties:

- Result of the algorithm does not depend on choices of $B \in \mathcal{B}$
  $\rightarrow$ denote the result as $P_{\mathcal{B}}(J)$
- $P_{\mathcal{B}}$ is connected to the complete lattice $\{J \mid J \text{ is } \mathcal{B}\text{-consistent}\} \cup \{\bot\}$
  $\rightarrow$ $P_{\mathcal{B}}$ is the maximal $\mathcal{B}$-consistent subset of $J$ (if $J$ has such a subset)
- If $P_{\mathcal{B}}(J) = \bot$:
  $\rightarrow$ original system is infeasible (but *not* vice versa)
  $\rightarrow$ it is possible to construct certificate of infeasibility $\bar{y}$

# Relation Between BCD and Constraint Propagation

**Theorem:** Let $y$ be a feasible point for dual (1), then:

- $y$ is an LM of dual (1) w.r.t. $\mathcal{B}$ if and only if $P_B(J(y)) \neq \bot \; \forall B \in \mathcal{B}$
- $y$ is an ILM of dual (1) w.r.t. $\mathcal{B}$ if and only if $J(y)$ is $\mathcal{B}$-consistent, i.e., $P_{\mathcal{B}}(J(y)) = J(y)$
- $y$ is a pre-ILM of dual (1) w.r.t. $\mathcal{B}$ if and only if $P_{\mathcal{B}}(J(y)) \neq \bot$.

**Corollary:** The following are equivalent:

- For all dual-feasible $y$, if complementary slackness (5) is infeasible, then $P_{\mathcal{B}}(J(y)) = \bot$ (refutation-completeness).
- Any (pre-)ILM $y$ of the dual (1) w.r.t. $\mathcal{B}$ is a global minimum.

These results can be generalized to any linear program in any form.

# BCD and Constraint Propagation in Existing Algorithms

For basic LP relaxation of Weighted CSP:

- VAC[1] / AugDAG[4,5] algorithm correspond to the primal-dual approach with the specified propagation rule
- Max-sum diffusion[6] satisfies[3] relative interior rule $\to$ fixed points are ILMs (i.e., points $y$ where $J(y)$ is $\mathcal{B}$-consistent)
- Fixed points of MPLP[7] and MPLP++[8] are[3] pre-ILMs ($P_{\mathcal{B}}(J(y)) \neq \perp$ means node-edge agreement, i.e., non-empty arc consistency closure)

---

[1] Cooper et al.: Soft arc consistency revisited

[4] Koval and Schlesinger: Two-dimensional Programming in Image Analysis Problems

[5] Werner: A Linear Programming Approach to Max-sum Problem: A Review

[6] Kovalevsky and Koval: A diffusion algorithm for decreasing energy of max-sum labeling problem

[3] Werner et al.: Relative interior rule in block-coordinate descent

[7] Globerson and Jaakkola: Fixing max-product: Convergent message passing algorithms for MAP LP-relaxations

[8] Tourani et al.: MPLP++: Fast, parallel dual block-coordinate ascent for dense graphical models

# BCD and Constraint Propagation in Existing Algorithms

LP relaxation of SAT (Boolean satisfiability problem):

- $P_{\mathcal{B}}$ performs unit propagation[9]
- Fixed points of BCD (with relative interior rule) on the dual (implicitly) define variables set to true/false via $J(y)$

[9] Dlask: Unit Propagation by Means of Coordinate-Wise Minimization