

A standard LP problem and its Lagrangian dual

$$\begin{aligned} v_{LP} = \text{minimum } & \mathbf{c}^T \mathbf{x}, \\ \text{subject to } & \mathbf{Ax} \leq \mathbf{b}, \\ & \mathbf{Dx} \leq \mathbf{d}, \\ & \mathbf{x} \in \mathbb{R}_+^n. \end{aligned}$$

- Let $X := \{\mathbf{x} \in \mathbb{R}_+^n \mid \mathbf{Ax} \leq \mathbf{b}\}$.
- We suppose for now that X is bounded.
- Further, let $P_X := \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^K\}$ be the set of extreme points in the polyhedron X .

Cutting Plane, Column generation and Dantzig–Wolfe decomposition

23 September 2008

0-0

- So,

$$\begin{aligned} v_L := \text{maximum } & z, \\ \text{subject to } & z \leq \mathbf{c}^T \mathbf{x}^i + \boldsymbol{\mu}^T (\mathbf{Dx}^i - \mathbf{d}), \quad i \in P_X, \\ & \boldsymbol{\mu} \geq \mathbf{0}. \end{aligned}$$

- We know that if at an optimal dual solution $\boldsymbol{\mu}^*$, the set $X(\boldsymbol{\mu}^*)$ is a singleton, then thanks to strong duality this solution is optimal (and it is unique!). This typically does not happen, unless an optimal solution \mathbf{x}^* happens to be an extreme point of X . We know, however, that \mathbf{x}^* always can be written as a convex combination of such points. Let's see how it can be generated.

- Its Lagrangian dual with respect to Lagrangian relaxing the constraints $\mathbf{Dx} \leq \mathbf{d}$ is to find

$$\begin{aligned} v_{LP} = v_L := \text{maximum } & q(\boldsymbol{\mu}), \\ \text{subject to } & \boldsymbol{\mu} \geq \mathbf{0}, \end{aligned}$$

where

$$\begin{aligned} q(\boldsymbol{\mu}) &:= \text{minimum}_{\mathbf{x} \in X} \{ \mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (\mathbf{Dx} - \mathbf{d}) \} \\ &= \text{minimum}_{i \in P_X} \{ \mathbf{c}^T \mathbf{x}^i + \boldsymbol{\mu}^T (\mathbf{Dx}^i - \mathbf{d}) \}. \end{aligned}$$

- Equivalent statement:

$$q(\boldsymbol{\mu}) \leq \mathbf{c}^T \mathbf{x}^i + \boldsymbol{\mu}^T (\mathbf{Dx}^i - \mathbf{d}), \quad i \in P_X, \quad \boldsymbol{\mu} \geq \mathbf{0}.$$

- Let $(\boldsymbol{\mu}^{k+1}, z^{k+1})$ be the solution to the above problem. If $z^{k+1} \leq \mathbf{c}^T \mathbf{x}^i + (\boldsymbol{\mu}^{k+1})^T (\mathbf{D}\mathbf{x}^i - \mathbf{d})$ holds for all $i \in P_X$, then $\boldsymbol{\mu}^{k+1}$ is optimal in the dual! Why?
- How to check optimality: find the most violated dual constraint! That is, solve the subproblem to find

$$q(\boldsymbol{\mu}^{k+1}) := \underset{\mathbf{x} \in X}{\text{minimum}} \{ \mathbf{c}^T \mathbf{x} + (\boldsymbol{\mu}^{k+1})^T (\mathbf{D}\mathbf{x} - \mathbf{d}) \} \quad (2)$$

$$= \underset{i \in P_X}{\text{minimum}} \{ \mathbf{c}^T \mathbf{x}^i + (\boldsymbol{\mu}^{k+1})^T (\mathbf{D}\mathbf{x}^i - \mathbf{d}) \}.$$

A cutting plane method for the Lagrangian dual problem

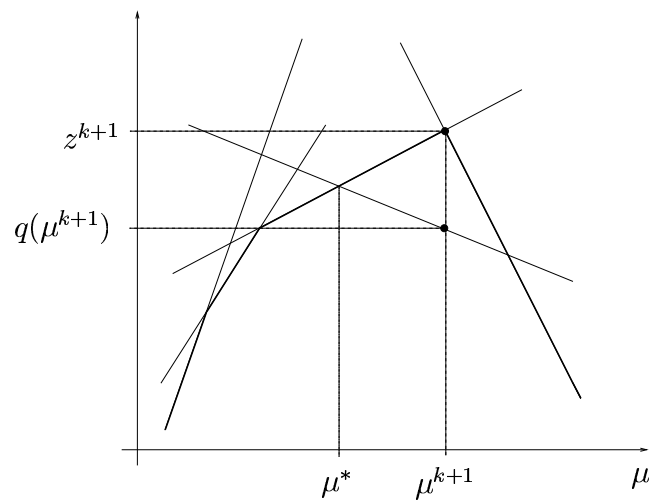
- Suppose only a subset of P_X is known, and consider the following restriction of the Lagrangian dual problem:

$$z^{k+1} := \max z, \quad (1a)$$

$$\text{s.t. } z \leq \mathbf{c}^T \mathbf{x}^i + \boldsymbol{\mu}^T (\mathbf{D}\mathbf{x}^i - \mathbf{d}), \quad i = 1, \dots, k, \quad (1b)$$

$$\boldsymbol{\mu} \geq \mathbf{0}. \quad (1c)$$

- How do we determine if we have found the optimal solution? And what IS the optimal solution when we find it?



- If $z^{k+1} \leq q(\boldsymbol{\mu}^{k+1})$ then $\boldsymbol{\mu}^{k+1}$ is optimal in the dual; otherwise, we have identified a constraint of the form $z \leq \mathbf{c}^T \mathbf{x}^i + \boldsymbol{\mu}^T (\mathbf{D}\mathbf{x}^i - \mathbf{d})$, where $i \in P_X$, which is violated at $(\boldsymbol{\mu}^{k+1}, z^{k+1})$. Add this inequality and re-solve the LP problem!
- We refer to this algorithm as a *cutting plane* algorithm, for the reason that it is based on adding constraints to the dual problem in order to improve the solution, in the process cutting off the previous point.
- Consider the figure on the next slide. The thick lines correspond to the subset of k inequalities known at iteration k .

Duality relationships and the Dantzig–Wolfe algorithm

- We rewrite the problem (1) as follows:

$$\begin{aligned} & \underset{(z, \boldsymbol{\mu})}{\text{maximize}} \quad z, \\ & \text{subject to} \quad z - \boldsymbol{\mu}^T(\mathbf{D}\mathbf{x}^i - \mathbf{d}) \leq \mathbf{c}^T\mathbf{x}^i, \quad i = 1, \dots, k, \\ & \quad \quad \quad \boldsymbol{\mu} \geq \mathbf{0}. \end{aligned}$$

- Obviously, $z^{k+1} \geq q(\boldsymbol{\mu}^{k+1})$ must hold, because of the possible lack of constraints. In this case, $z^{k+1} > q(\boldsymbol{\mu}^{k+1})$ holds, so in the next step when we evaluate $q(\boldsymbol{\mu}^{k+1})$ we can identify and add the last lacking inequality; the resulting maximization will then yield the optimal solution $\boldsymbol{\mu}^*$ shown in the picture.
- How do we generate a primal optimal solution from this scheme? Let us look at the dual of the problem (1) in this cutting plane algorithm.

$$v^{k+1} = \text{minimum} \quad \mathbf{c}^T \left(\sum_{i=1}^k \lambda_i \mathbf{x}^i \right), \quad (3)$$

$$\begin{aligned} \text{subject to} \quad & \sum_{i=1}^k \lambda_i = 1, \\ & \lambda_i \geq 0, \quad i = 1, \dots, k, \end{aligned}$$

$$\mathbf{D} \left(\sum_{i=1}^k \lambda_i \mathbf{x}^i \right) \leq \mathbf{d}.$$

- We maximize $\mathbf{c}^T \mathbf{x}$ subject to \mathbf{x} lying in the convex hull of the extreme points \mathbf{x}^i found so far *and* fulfilling the constraints that are Lagrangian relaxed.

- With LP dual variables $\lambda_i \geq 0$ for the linear constraints, we obtain the LP dual to find

$$v^{k+1} = \text{minimum} \quad \sum_{i=1}^k (\mathbf{c}^T \mathbf{x}^i) \lambda_i,$$

$$\text{subject to} \quad \sum_{i=1}^k \lambda_i = 1,$$

$$-\sum_{i=1}^k (\mathbf{D}\mathbf{x}^i - \mathbf{d}) \lambda_i \geq \mathbf{0},$$

$$\lambda_i \geq 0, \quad i = 1, \dots, k,$$

that is,

- Three algorithms which are “dual” to each other:

Cutting plane applied to the Lagrangian dual

\iff

Dantzig–Wolfe applied to the original LP

\iff

Benders decomposition applied to the dual LP.

- The problem (3) is known as the *restricted master problem* (RMP) in the Dantzig–Wolfe algorithm.
- In this algorithm, we have at hand a subset $\{1, \dots, k\}$ of extreme points of X (and a dual vector $\boldsymbol{\mu}^k$), and find a feasible solution to the original LP problem by solving the restricted master problem (3). We then generate an optimal dual solution $\boldsymbol{\mu}^{k+1}$ to this restricted problem, corresponding to the constraints $\mathbf{D}\mathbf{x} \leq \mathbf{d}$. If and only if the vector \mathbf{x}^i generated in the next subproblem (2) was already included, we have found the optimal solution to the problem.

Basic feasible solutions

$B = \{m \text{ elements from the set } \{1, \dots, n\}\}$ is a basis if the corresponding matrix $\mathbf{B} = (\mathbf{a}_j)_{j \in B}$ has an inverse, \mathbf{B}^{-1}

A basic solution is given by $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$ and $x_j = 0, j \notin B$.

It is feasible if $\mathbf{x}_B \geq \mathbf{0}^m$

A better basic feasible solution can be found by computing reduced costs: $\bar{c}_j = c_j - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{a}_j$ for $j \notin B$

Let $\bar{c}_s = \text{minimum}_{j \notin B} \bar{c}_j$

If $\bar{c}_s < 0 \implies$ a better solution is received if x_s enters the basis

If $\bar{c}_s \geq 0 \implies \mathbf{x}_B$ is an optimal basic solution

Column generation

An LP with very many variables $c_j, x_j \in \mathbb{R}, \mathbf{a}_j, \mathbf{b} \in \mathbb{R}^m, m \ll n$

$$\begin{aligned} \text{minimize } z &= \sum_{j=1}^n c_j x_j \\ \text{subject to } & \sum_{j=1}^n \mathbf{a}_j x_j = \mathbf{b} \\ & x_j \geq 0, \quad j = 1, \dots, n \end{aligned}$$

The matrix $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ is too large to handle. Assume that m is relatively small \implies the basic matrix is not too large ($m \times m$)

Example: The Cutting Stock Problem

Supply: rolls of e.g. paper of length L

Demand: b_i pieces of length $\ell_i < L$, $i = 1, \dots, m$

Objective: minimize the number of rolls needed to satisfy the demanded of the pieces

First formulation:

Let

$$x_k = \begin{cases} 1 & \text{roll } k \text{ is used} \\ 0 & \text{otherwise} \end{cases} \quad y_{ik} = \begin{cases} 1 & \text{piece } i \text{ is cut from roll } k \\ 0 & \text{otherwise} \end{cases}$$

Suppose the columns \mathbf{a}_j are defined by a set $S = \{\mathbf{a}_j \mid j = 1, \dots, n\}$ being, e.g., solutions to a system of equations (extreme points, integer points, ...)

The incoming column is then chosen by solving a “subproblem”:

$$\bar{c}(\mathbf{a}') = \underset{\mathbf{a}' \in S}{\text{minimum}} \{c - \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{a}'\}$$

\mathbf{a}' is a column having the least reduced cost wrt basis B

If $\bar{c}(\mathbf{a}') < 0$ let the column $\begin{pmatrix} c(\mathbf{a}') \\ \mathbf{a}' \end{pmatrix}$ enter problem

Second formulation:

Cut pattern j contains a_{ij} pieces of length ℓ_i

Feasible pattern if $\sum_{i=1}^m \ell_i a_{ij} \leq L$, where $a_{ij} \geq 0$, integer

Integer variables: x_j = number of times pattern j is used

Bad news: n = total number of feasible cut pattern — very large integer

$$\begin{aligned} &\text{minimize} && \sum_{k=1}^M x_k \\ &\text{s.t.} && \sum_{i=1}^m \ell_i y_{ik} \leq L x_k, \quad k = 1, \dots, M \\ &&& \sum_{k=1}^M y_{ij} = b_i, \quad i = 1, \dots, m \\ &&& x_k, y_{ik} \geq 0, \text{ binary,} \end{aligned}$$

The value of the LP-relaxation is $\frac{\sum \ell_i b_i}{L}$ which can be very bad if $\ell_i = \lfloor L/2 + 1 \rfloor$ for large L (large duality gap, potentially bad performance of ILP-solvers).

Start solution

Trivial: m unit columns (gives lots of waste):

$$\begin{aligned} & \text{minimize} && \sum_{j=1}^m x_j \\ & \text{subject to} && x_j = b_j, \quad j = 1, \dots, m \\ & && x_j \geq 0, \quad j = 1, \dots, m \end{aligned}$$

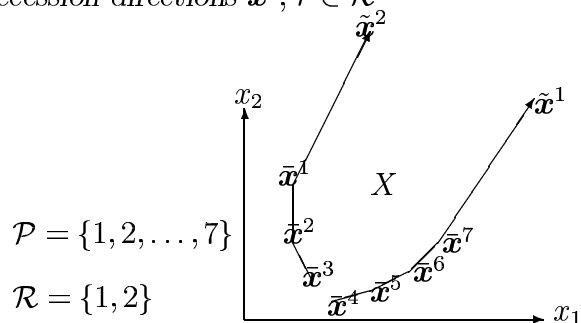
$$\begin{aligned} & \text{minimize} && \sum_{j=1}^n x_j \\ & \text{subject to} && \sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, \dots, m \\ & && x_j \geq 0, \text{ integer}, \quad j = 1, \dots, n \end{aligned}$$

Good news: the value of the LP-relaxation is often very close to the value of the optimal solution^a. We may relax the integrality constraints and solve an LP instead of an ILP!

^aMarcotte 1985: The cutting stock problem and integer rounding, Mathematical Programming 33

Formulation of LP on column generation form—Dantzig–Wolfe decomposition

Let $X = \{\mathbf{x} \in \mathbb{R}_+^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}\}$ (or $\mathbf{A}\mathbf{x} \leq \mathbf{b}$) be a polyhedron with the *extreme points* $\bar{\mathbf{x}}^p, p \in \mathcal{P}$ and the *extreme recession directions* $\tilde{\mathbf{x}}^r, r \in \mathcal{R}$



New columns

Generate better patterns using the dual variables π :

$$\begin{aligned} & 1 + \text{maximum}_{a_{ij}} \sum_{i=1}^m \pi_i a_{ij} && \left[\text{minimize} (c_j - \underbrace{\mathbf{c}_B^T \mathbf{B}^{-1}}_{\pi} \mathbf{a}_j) \right] \\ & \text{subject to} && \sum_{i=1}^m \ell_i a_{ij} \leq L, \\ & && a_{ij} \geq 0, \text{ integer}, \quad i = 1, \dots, m \end{aligned}$$

Solution to this knapsack problem: New column \mathbf{a}_j

An LP and its complete master problem

$$\begin{aligned}
 \text{[LP1]} \quad z^* = \text{minimum } \mathbf{c}^T \mathbf{x} \\
 \text{subject to } \mathbf{Ax} = \mathbf{b} \text{ ("simple" constraints)} \\
 \mathbf{Dx} = \mathbf{d} \text{ (complicating constraints)} \\
 \mathbf{x} \geq \mathbf{0}
 \end{aligned}$$

Let $X = \{\mathbf{x} \geq 0 \mid \mathbf{Ax} = \mathbf{b}\}$ with the extreme points $\tilde{\mathbf{x}}^p$, $p \in \mathcal{P}$ and the extreme directions $\tilde{\mathbf{x}}^r$, $r \in \mathcal{R} \implies$

$$\mathbf{x} \in X \iff \left(\begin{array}{l} \mathbf{x} = \sum_{p \in \mathcal{P}} \lambda_p \tilde{\mathbf{x}}^p + \sum_{r \in \mathcal{R}} \mu_r \tilde{\mathbf{x}}^r \\ \sum_{p \in \mathcal{P}} \lambda_p = 1 \\ \lambda_p \geq 0, \quad p \in \mathcal{P} \\ \mu_r \geq 0, \quad r \in \mathcal{R} \end{array} \right)$$

$\mathbf{x} \in X$ is a convex combination of the extreme points plus a conical combination of the extreme directions

This *inner representation* of the set X can be used to reformulate a linear optimization problem according to the *Dantzig-Wolfe decomposition principle*, which is then solved by column generation.

The dual of [LP2] is given by (not all extreme pts./dirs. found yet: $\bar{\mathcal{P}} \subseteq \mathcal{P}$; $\bar{\mathcal{R}} \subseteq \mathcal{R}$)

$$\begin{aligned}
 \text{[DLP2]} \quad z^* \leq \max_{(\boldsymbol{\pi}, q)} \mathbf{d}^T \boldsymbol{\pi} + q \\
 \text{s.t. } (\mathbf{D}\tilde{\mathbf{x}}^p)^T \boldsymbol{\pi} + q \leq (\mathbf{c}^T \tilde{\mathbf{x}}^p), \quad p \in \bar{\mathcal{P}} \quad | \quad \lambda_p \\
 (\mathbf{D}\tilde{\mathbf{x}}^r)^T \boldsymbol{\pi} \leq (\mathbf{c}^T \tilde{\mathbf{x}}^r), \quad r \in \bar{\mathcal{R}} \quad | \quad \mu_r
 \end{aligned}$$

with solutions $(\bar{\boldsymbol{\pi}}, \bar{q})$

Reduced cost for the variable λ_p , $p \in \mathcal{P} \setminus \bar{\mathcal{P}}$ is given by $(\mathbf{c}^T \tilde{\mathbf{x}}^p) - (\mathbf{D}\tilde{\mathbf{x}}^p)^T \bar{\boldsymbol{\pi}} - \bar{q} = (\mathbf{c} - \mathbf{D}^T \bar{\boldsymbol{\pi}})^T \tilde{\mathbf{x}}^p - \bar{q}$

Reduced cost for the variable μ_r , $r \in \mathcal{R} \setminus \bar{\mathcal{R}}$ is given by $(\mathbf{c}^T \tilde{\mathbf{x}}^r) - (\mathbf{D}\tilde{\mathbf{x}}^r)^T \bar{\boldsymbol{\pi}} = (\mathbf{c} - \mathbf{D}^T \bar{\boldsymbol{\pi}})^T \tilde{\mathbf{x}}^r$

$$\begin{aligned}
 \text{[LP2]} \quad z^* = \min \sum_{p \in \mathcal{P}} \lambda_p (\mathbf{c}^T \tilde{\mathbf{x}}^p) + \sum_{r \in \mathcal{R}} \mu_r (\mathbf{c}^T \tilde{\mathbf{x}}^r) \\
 \text{s.t. } \sum_{p \in \mathcal{P}} \lambda_p (\mathbf{D}\tilde{\mathbf{x}}^p) + \sum_{r \in \mathcal{R}} \mu_r (\mathbf{D}\tilde{\mathbf{x}}^r) = \mathbf{d} \quad | \quad \boldsymbol{\pi} \\
 \sum_{p \in \mathcal{P}} \lambda_p = 1 \quad | \quad q \\
 \lambda_p, \mu_r \geq 0, \quad \forall p, r
 \end{aligned}$$

Number of constraints in [LP2] equals to "the number of constraints in $\mathbf{Dx} = \mathbf{d}$ " + 1

Number of columns very large (# extreme pts./dirs. to X)

Column generation

The least reduced cost is found by solving the subproblem

$$\min_{x \in X} (\mathbf{c} - \mathbf{D}^T \boldsymbol{\pi})^T \mathbf{x} \quad \left(\text{alt: } \min_{x \in X} (\mathbf{c} - \mathbf{D}^T \bar{\boldsymbol{\pi}})^T \mathbf{x} - \bar{q} \right)$$

Gives as solution an extreme point, $\bar{\mathbf{x}}^p$, or an extreme direction $\tilde{\mathbf{x}}^r$

\implies a new column in [LP2]: (if < 0)

Either $\begin{pmatrix} \mathbf{c}^T \bar{\mathbf{x}}^p \\ \mathbf{D} \bar{\mathbf{x}}^p \\ 1 \end{pmatrix}$ or $\begin{pmatrix} \mathbf{c}^T \tilde{\mathbf{x}}^r \\ \mathbf{D} \tilde{\mathbf{x}}^r \\ 0 \end{pmatrix}$ enters the problem and improves the solution