Lecture 8–10: Column generation, Dantzig–Wolfe decomposition, Cutting plane methods, Benders decomposition, and Branch–and–price—Not much is new under the sun

Ann-Brith Strömberg
Michael Patriksson

19 April 2004

A standard LP problem and its Lagrangian dual

$$
\begin{aligned}
v_{LP} = \text{minimum} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\
\text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}, \\
& \boldsymbol{D}\boldsymbol{x} \le \boldsymbol{d}, \\
& \boldsymbol{x} \in \mathbb{R}^n_+.
\end{aligned}
$$

- We suppose for now that X is bounded.
- Let $P_X := \{\boldsymbol{x}^1, \boldsymbol{x}^2, \dots, \boldsymbol{x}^K\}$ be the set of extreme points in the polyhedron $X := \{\, \boldsymbol{x} \in \mathbb{R}^n_+ \mid \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \,\}$.

- Its Lagrangian dual with respect to Lagrangian relaxing the constraints $\boldsymbol{D}\boldsymbol{x} \le \boldsymbol{d}$ is to find

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

where

$$
\begin{aligned}
q(\boldsymbol{\mu}) &:= \underset{\boldsymbol{x} \in X}{\text{minimum}} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x} - \boldsymbol{d}) \right\} \\
&= \underset{i \in P_X}{\text{minimum}} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}) \right\}.
\end{aligned}
$$

- Equivalent statement:

$$
q(\boldsymbol{\mu}) \le \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}), \qquad i \in P_X, \quad \boldsymbol{\mu} \ge \boldsymbol{0}.
$$

- So,

$$
\begin{aligned}
v_L := \text{maximum} \quad & z, \\
\text{subject to} \quad & z \le \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}), \qquad i \in P_X, \\
& \boldsymbol{\mu} \ge \boldsymbol{0}.
\end{aligned}
$$

- 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 $\boldsymbol{x}^*$ happens to be an extreme point of X. We know, however, that $\boldsymbol{x}^*$ always can be written as a convex combination of such points. Let's see how it can be generated.

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:

$$z^{k+1} := \max \ z, \tag{1a}$$
$$\text{s.t.} \ z \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}), \quad i = 1, \ldots, k, \tag{1b}$$
$$\boldsymbol{\mu} \geq \boldsymbol{0}. \tag{1c}$$

- How do we determine if we have found the optimal solution? And what IS the optimal solution when we find it?
- Let $(\boldsymbol{\mu}^{k+1}, z^{k+1})$ be the solution to the above problem.

If $z^{k+1} \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + (\boldsymbol{\mu}^{k+1})^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d})$ holds for all $i \in P_X$, then $\boldsymbol{\mu}^{k+1}$ is optimal in the dual! Why?

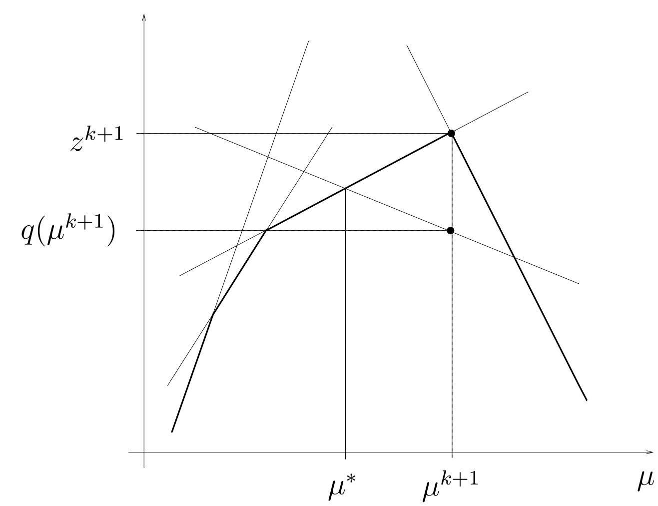
- How to check optimality: find the most violated dual constraint! That is, solve the subproblem to find

$$q(\boldsymbol{\mu}^{k+1}) := \underset{\boldsymbol{x} \in X}{\text{minimum}} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + (\boldsymbol{\mu}^{k+1})^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x} - \boldsymbol{d}) \right\} \tag{2}$$
$$= \underset{i \in P_X}{\text{minimum}} \left\{ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + (\boldsymbol{\mu}^{k+1})^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d}) \right\}.$$

- 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 $z \leq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d})$, 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 below picture. The think lines correspond to the subset of k inequalities known at iteration k.



- Obviously, $z^{k+1} \geq q(\mu^{k+1})$ must hold, because of the possible lack of constraints. In this case, $z^{k+1} > q(\mu^{k+1})$ holds, so in the next step when we evaluate $q(\mu^{k+1})$ we can identify and add the last lacking inequality; the resulting maximization will then yield the optimal solution μ^* shown in the picture.
- What is the relationship to the standard simplex method?
- 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.

Duality relationships and the Dantzig–Wolfe algorithm

- We rewrite the problem (1) as follows:

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

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

$$\begin{aligned} v^{k+1} = \text{minimum } & \sum_{i=1}^{k} (\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^i)\lambda_i, \\ \text{subject to } & \sum_{i=1}^{k} \lambda_i = 1, \\ & -\sum_{i=1}^{k} (\boldsymbol{D}\boldsymbol{x}^i - \boldsymbol{d})\lambda_i \geq \boldsymbol{0}, \\ & \lambda_i \geq 0, \qquad i = 1, \dots, k, \end{aligned}$$

that is,

$$\begin{aligned} v^{k+1} = \text{minimum } & \boldsymbol{c}^{\mathrm{T}} \left(\sum_{i=1}^{k} \lambda_i \boldsymbol{x}^i \right), \qquad (3) \\ \text{subject to } & \sum_{i=1}^{k} \lambda_i = 1, \\ & \lambda_i \geq 0, \qquad i = 1, \dots, k, \\ & \boldsymbol{D} \left(\sum_{i=1}^{k} \lambda_i \boldsymbol{x}^i \right) \leq \boldsymbol{d}. \end{aligned}$$

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

- 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 $\{1, \ldots, k\}$ 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 problem, corresponding to the constraints $\boldsymbol{Dx} \leq \boldsymbol{d}$. If and only if the vector $\boldsymbol{x}^i$ generated in the next subproblem (2) was already included, we have found the optimal solution to the problem.

- Three algorithms which are "dual" to each other:

Cutting plane applied to the Lagrangian dual

$\Longleftrightarrow$

Dantzig–Wolfe applied to the original LP

$\Longleftrightarrow$

Benders decomposition applied to the dual LP.

Column generation

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

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

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

Basic feasible solutions

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

A basic solution is given by $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b}$ and $x_j = 0$, $j \notin B$. It is feasible if $\boldsymbol{x}_B \geq \boldsymbol{0}^m$

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

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

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

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

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

The incoming column is then chosen by solving a "subproblem":

$$\text{Let } c(\boldsymbol{a}_j) = c_j;$$
$$\bar{c}(\boldsymbol{a}(B)) = \underset{\boldsymbol{a} \in S}{\text{minimum}} \left\{ c(\boldsymbol{a}) - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{a} \right\}$$

$\boldsymbol{a}(B)$ is a column having the least reduced cost w.r.t. the basis B

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

Example: Cutting stock

Supply: (long) pieces of wood of length L
Demand: b_i pieces of wood of length $\ell_i < L$, $i = 1, \dots, m$
Objective: minimize the number of pieces needed for producing the pieces demanded
Cut pattern: number 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
Variables: x_j = number of times pattern j is used
n = total number of feasible cut pattern — very large integer

Problem:

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

Start solution and new columns

Trivial: m unit columns (gives lots of waste) $\Longrightarrow$

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

Generate better patterns (integer knapsack problem): $\Longrightarrow$ new column

$$1 - \underset{a_{ik}}{\text{maximum}} \sum_{i=1}^{m} a_{ik} \qquad \left[\text{minimize}\,(c_k - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{a}_k)\right]$$
$$\text{subject to} \sum_{i=1}^{m} \ell_i a_{ik} \leq L,$$
$$a_{ik} \geq 0, \text{ integer}, \qquad i = 1, \ldots, m$$

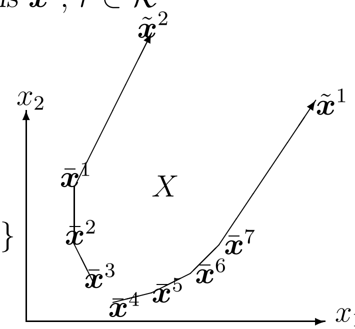
Solution: $\boldsymbol{a}_k$

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

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

$\mathcal{P} = \{1, 2, \ldots, 7\}$

$\mathcal{R} = \{1, 2\}$



$$\boldsymbol{x} \in X \Longleftrightarrow \left(\begin{array}{rcl} \boldsymbol{x} & = & \sum_{p \in \mathcal{P}} \lambda_p \bar{\boldsymbol{x}}^p + \sum_{r \in \mathcal{R}} \mu_r \tilde{\boldsymbol{x}}^r \\ & & \sum_{p \in \mathcal{P}} \lambda_p = 1 \\ & & \lambda_p \geq 0, \quad p \in \mathcal{P} \\ & & \mu_r \geq 0, \quad r \in \mathcal{R} \end{array} \right)$$

$\boldsymbol{x} \in X$ 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.

An LP and its complete master problem

$$\text{[LP1]} \quad z^* = \text{minimum } \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$$
$$\text{subject to } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \text{ (“simple” constraints)}$$
$$\boldsymbol{D}\boldsymbol{x} = \boldsymbol{d} \text{ (complicating constraints)}$$
$$\boldsymbol{x} \geq \boldsymbol{0}$$

Let $X = \{\, \boldsymbol{x} \geq 0 \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \,\}$ with the extreme points $\bar{\boldsymbol{x}}^p$, $p \in \mathcal{P}$ and the extreme directions $\tilde{\boldsymbol{x}}^r$, $r \in \mathcal{R} \Longrightarrow$

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

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

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

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

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

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
$$(\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^p) - (\boldsymbol{D}\bar{\boldsymbol{x}}^p)^{\mathrm{T}}\bar{\boldsymbol{\pi}} - \bar{q} = (\boldsymbol{c} - \boldsymbol{D}^{\mathrm{T}}\bar{\boldsymbol{\pi}})^{\mathrm{T}}\bar{\boldsymbol{x}}^p - \bar{q}$$

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

Column generation

The least reduced cost is found by solving the subproblem

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

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

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

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

Example

$$\begin{array}{llll} z^*_{\mathrm{IP}} & = \min & 2x_1 + 3x_2 + x_3 + 4x_4 & \\ \text{[IP]} & \text{s.t.} & 3x_1 + 2x_2 + 3x_3 + 2x_4 = 5 & |\ \boldsymbol{D}\boldsymbol{x} = \boldsymbol{d} \\ & & x_1 + x_2 + x_3 + x_4 = 2 & \\ & & x_1 \quad x_2 \quad x_3 \quad x_4 \in \{0,1\} & \end{array}$$

$$X = \left\{ \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\end{pmatrix} \right\} = \{\bar{\boldsymbol{x}}^1, \ldots, \bar{\boldsymbol{x}}^6\}$$

Optimal solution: $\boldsymbol{x}^*_{\mathrm{IP}} = (0,1,1,0)^{\mathrm{T}}$ $\quad z^*_{\mathrm{IP}} = 4$

LP-relaxation

$$\begin{array}{rll} z^* = & \min\ 2x_1+3x_2+x_3+4x_4 & [\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}] \\ \text{[LP1]} & \text{s.t. } 3x_1+2x_2+3x_3+2x_4=5 & [\boldsymbol{D}\boldsymbol{x}=\boldsymbol{d}] \\ & x_1+x_2+x_3+x_4=2 & [\boldsymbol{x}\in X] \\ & 0\le x_1 \quad x_2 \quad x_3 \quad x_4 \le 1 & [\boldsymbol{x}\in X] \end{array}$$

$$X = \operatorname{conv}\left\{ \begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\1\end{pmatrix}, \begin{pmatrix}0\\0\\1\\1\end{pmatrix} \right\} = \operatorname{conv}\{\bar{\boldsymbol{x}}^1,\dots,\bar{\boldsymbol{x}}^6\}$$

$$= \left\{ \boldsymbol{x}\in\mathbb{R}^4 \,\middle|\, \boldsymbol{x}=\textstyle\sum_{p=1}^6 \lambda_p\bar{\boldsymbol{x}}^p;\ \sum_{p=1}^6\lambda_p=1;\ \lambda_p\ge 0,\ p=1,\dots,6 \right\}$$

$$\begin{array}{ll} \text{[LP2]} \quad z^* = \min & 5\lambda_1+3\lambda_2+6\lambda_3+4\lambda_4+7\lambda_5+5\lambda_6 \\ \text{s.t.} & 5\lambda_1+6\lambda_2+5\lambda_3+5\lambda_4+4\lambda_5+5\lambda_6=5 \\ & \lambda_1+\lambda_2+\lambda_3+\lambda_4+\lambda_5+\lambda_6=1 \\ & \lambda_1,\ \lambda_2,\ \lambda_3,\ \lambda_4,\ \lambda_5,\ \lambda_6\ge 0 \end{array}$$

Start columns: $\lambda_1,\lambda_2,\lambda_3$

[LP2]

$$\begin{array}{ll} z^*\le\min & 5\lambda_1+3\lambda_2+6\lambda_3 \\ \text{s.t.} & 5\lambda_1+6\lambda_2+5\lambda_3=5 \\ & \lambda_1+\lambda_2+\lambda_3=1 \\ & \lambda_1,\ \lambda_2,\ \lambda_3\ge 0 \end{array}$$

[DLP2]

$$\begin{array}{ll} z^*\le\max & 5\pi+q \\ \text{s.t.} & 5\pi+q\le 5 \\ & 6\pi+q\le 3 \\ & 5\pi+q\le 6 \end{array}$$

Solution: $\bar{\boldsymbol{\lambda}}=(1,0,0)^{\mathrm{T}}$, $\bar{\pi}=-2$, $\bar{q}=15$

Reduced costs

$$\begin{aligned} &\min_{\boldsymbol{x}\in X}(\boldsymbol{c}-\boldsymbol{D}^{\mathrm{T}}\bar{\pi})^{\mathrm{T}}\boldsymbol{x}-\bar{q} \\ =\ &\min_{p=1,\dots,6}(\boldsymbol{c}-\boldsymbol{D}^{\mathrm{T}}\bar{\pi})^{\mathrm{T}}\bar{\boldsymbol{x}}^p-\bar{q} \\ =\ &\min_{p=1,\dots,6}\{[(2,3,1,4)-(3,2,3,2)\cdot(-2)]\,\bar{\boldsymbol{x}}^p-15\} \\ =\ &\min\{0,0,1,-1,0,0\}=-1<0 \end{aligned}$$

New extreme point in [LP1]: $\bar{\boldsymbol{x}}^4=(0,1,1,0)^{\mathrm{T}}$

Column in [LP2]: $\begin{pmatrix}\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^4\\ \boldsymbol{A}\bar{\boldsymbol{x}}^4\\ 1\end{pmatrix}=\begin{pmatrix}4\\5\\1\end{pmatrix}$

New, extended problem

[LP2]

$$\begin{array}{ll} z^*\le\min & 5\lambda_1+3\lambda_2+6\lambda_3+4\lambda_4 \\ \text{s.t.} & 5\lambda_1+6\lambda_2+5\lambda_3+5\lambda_4=5 \\ & \lambda_1+\lambda_2+\lambda_3+\lambda_4=1 \\ & \lambda_1,\ \lambda_2,\ \lambda_3,\ \lambda_4\ge 0 \end{array}$$

[DLP2]

$$\begin{array}{ll} z^*\le\max & 5\pi+q \\ \text{s.t.} & 5\pi+q\le 5 \\ & 6\pi+q\le 3 \\ & 5\pi+q\le 6 \\ & 5\pi+q\le 4 \end{array}$$

Solution:

$\bar{\boldsymbol{\lambda}}=(0,0,0,1)^{\mathrm{T}}$, $\bar{\pi}=-1$, $\bar{q}=9$

Reduced costs:

$\min_{p=1,\dots,6}\{(5,5,4,6)\,\bar{\boldsymbol{x}}^p-9\}=\min\{1,0,2,0,2,1\}=0$

Optimal solution to [LP2] and [LP1]

$\boldsymbol{\lambda}^* = (0,0,0,1,0,0)^{\mathrm{T}}, \qquad \pi^* = -1, \quad q^* = 9$

$\Longrightarrow \boldsymbol{x}^* = \bar{\boldsymbol{x}}^4 = (0,1,1,0)^{\mathrm{T}} = \boldsymbol{x}^*_{\mathrm{IP}}, \qquad z^* = 4 = z^*_{\mathrm{IP}}$

It was a coincidence that the solution was integral!

In general, the solution $\boldsymbol{x}^*$ to [LP1] can have fractional variable values.

Solution to [IP]

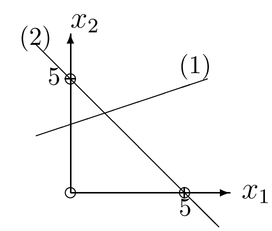
We need to find an integral solution (not certainly an optimal solution to [IP]) among the columns generated, i.e., solve

$$\min\left\{(2,3,1,4)\boldsymbol{x} \,\middle|\, (3,2,3,2)\boldsymbol{x} = 5,\ \boldsymbol{x} \in \{\bar{\boldsymbol{x}}^1, \bar{\boldsymbol{x}}^2, \bar{\boldsymbol{x}}^3, \bar{\boldsymbol{x}}^4\}\right\}$$

Numerical example of Dantzig-Wolfe decomposition

$$\begin{array}{llcrcl}
\min & x_1 & - & 3x_2 & & & (0) \\
\text{då} & -x_1 & + & 2x_2 & \leq & 6 & (1) \quad \text{(complicating)} \\
 & x_1 & + & x_2 & \leq & 5 & (2) \\
 & x_1 & , & x_2 & \geq & 0 & (3)
\end{array}$$

$$\begin{aligned} X &= \left\{\boldsymbol{x} \in \mathbb{R}^2_+ \,\middle|\, x_1 + x_2 \leq 5\right\} \\ &= \operatorname{conv}\left\{(0,0)^{\mathrm{T}}, (0,5)^{\mathrm{T}}, (5,0)^{\mathrm{T}}\right\} \end{aligned}$$



Complete DW-master problem

$$\boldsymbol{x} \in X \iff \begin{cases} \boldsymbol{x} = \lambda_1 \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 5 \end{pmatrix} + \lambda_3 \begin{pmatrix} 5 \\ 0 \end{pmatrix} = \begin{pmatrix} 5\lambda_3 \\ 5\lambda_2 \end{pmatrix} \\ \lambda_1 + \lambda_2 + \lambda_3 = 1 \\ \lambda_1, \lambda_2, \lambda_3 \geq 0 \end{cases}$$

$$\begin{array}{lll}
\min & -15\lambda_2 + 5\lambda_3 & (0) \\
\text{s.t.} & 10\lambda_2 - 5\lambda_3 \leq 6 & (1) \\
 & \lambda_1 + \lambda_2 + \lambda_3 = 1 & \\
 & \lambda_1, \lambda_2, \lambda_3 \geq 0 &
\end{array}$$

The first master problem is constructed from the points $(0,0)^{\mathrm{T}}$ and $(0,5)^{\mathrm{T}}$ (corresponds to λ_1 and λ_2)

Iteration 1

$$\begin{array}{lll}
\min & -15\lambda_2 & (0) \\
\text{s.t.} & 10\lambda_2 \leq 6 & (1) \\
 & \lambda_1 + \lambda_2 = 1 & \\
 & \lambda_1, \lambda_2 \geq 0 &
\end{array}$$

Solution: $\boldsymbol{\lambda} = (\frac{2}{5}, \frac{3}{5})^{\mathrm{T}}$

Dual solution: $\pi = -\frac{3}{2}, q = 0$

Least reduced cost: $\min_{\boldsymbol{x} \in X} \left[(\boldsymbol{c}^{\mathrm{T}} - \pi \boldsymbol{D})\boldsymbol{x} - q\right]$

$= \min_{\boldsymbol{x} \in X} \left(\left[(1,-3) - (-\frac{3}{2})(-1,2)\right] \boldsymbol{x} - 0\right)$

$= \min\left\{-\frac{1}{2}x_1 \mid x_1 + x_2 \leq 5; \boldsymbol{x} \geq \boldsymbol{0}^2\right\} = -\frac{5}{2} < 0 \Longrightarrow \bar{\boldsymbol{x}} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$

New column: $\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}} = (1,-3)(5,0)^{\mathrm{T}} = 5$

$\boldsymbol{D}\bar{\boldsymbol{x}} = (-1,2)(5,0)^{\mathrm{T}} = -5$

$\Longrightarrow \begin{pmatrix} \frac{5}{-5} \\ 1 \end{pmatrix}$

Iteration 2

$$\begin{array}{lll} \min & -15\lambda_2 + 5\lambda_3 & \\ \text{s.t.} & 10\lambda_2 - 5\lambda_3 \le 6 & \text{Solution: } \boldsymbol{\lambda} = (0, \tfrac{11}{15}, \tfrac{4}{15})^{\mathrm{T}} \\ & \lambda_1 + \lambda_2 + \lambda_3 = 1 & \text{Dual solution: } \pi = -\tfrac{4}{3}, q = -\tfrac{5}{3} \\ & \lambda_1, \lambda_2, \lambda_3 \ge 0 & \end{array}$$

Least reduced cost: $\min_{\boldsymbol{x} \in X} \left[(\boldsymbol{c}^{\mathrm{T}} - \pi \boldsymbol{D})\boldsymbol{x} - q \right]$

$= \min_{\boldsymbol{x} \in X} \left(\left[(1, -3) - (-\tfrac{4}{3})(-1, 2) \right] \boldsymbol{x} - (-\tfrac{5}{3}) \right)$

$= \min \{ -\tfrac{1}{3}x_1 - \tfrac{1}{3}x_2 + \tfrac{5}{3} \mid x_1 + x_2 \le 5; \boldsymbol{x} \ge \boldsymbol{0}^2 \} = 0$

Optimal solution: $\boldsymbol{\lambda}^* = (0, \tfrac{11}{15}, \tfrac{4}{15})^{\mathrm{T}}$

$\Longrightarrow \boldsymbol{x}^* = (5\lambda_3, 5\lambda_2)^{\mathrm{T}} = (\tfrac{4}{3}, \tfrac{11}{3})^{\mathrm{T}}; \quad z^* = \tfrac{4}{3} - 3 \cdot \tfrac{11}{3} = -9\tfrac{2}{3}$

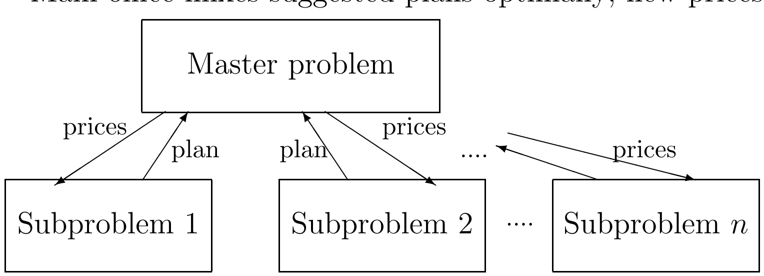
Block-angular structure

$$\begin{array}{lllll} \max & \boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1 + & \boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_2 + \cdots + & \boldsymbol{c}_n^{\mathrm{T}}\boldsymbol{x}_n & \\ \text{s.t.} & \boldsymbol{D}_1\boldsymbol{x}_1 + & \boldsymbol{D}_2\boldsymbol{x}_2 + \cdots + & \boldsymbol{D}_n\boldsymbol{x}_n \le \boldsymbol{d} & \mid \boldsymbol{\pi} \\ & \boldsymbol{A}_1\boldsymbol{x}_1 & & \le \boldsymbol{b}_1 & \mid \boldsymbol{x}_1 \in X_1 \\ & & \boldsymbol{A}_2\boldsymbol{x}_2 & \le \boldsymbol{b}_2 & \mid \boldsymbol{x}_2 \in X_2 \\ & & \ldots & \ddots & \\ & & & \boldsymbol{A}_n\boldsymbol{x}_n \le \boldsymbol{b}_n & \mid \boldsymbol{x}_n \in X_n \\ & & \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n & \ge \boldsymbol{0} & \end{array}$$

$$X = X_1 \times X_2 \times \ldots \times X_n$$

DW decomposition as decentralized planning

- Main office (master problem) sets prizes ($\boldsymbol{\pi}$) for the common resources (complicating constraints).
- Departments (subproblems) suggest (production) plans ($\boldsymbol{D}_j \bar{\boldsymbol{x}}_j^p$) based on given prices.
- Main office mixes suggested plans optimally; new prices.



Find feasible solutions (right-hand side allocation)

Let $\bar{\lambda}_p^j$, $p \in \mathcal{P}$, and $\bar{\mu}_r^j$, $r \in \mathcal{R}$, $j = 1, \ldots, n$, be a feasible and (almost) optimal solution to the master problem. A good feasible $\boldsymbol{x}$-solution is then given by (for all j):

$$\begin{array}{ll} \text{maximize} & \boldsymbol{c}_j^{\mathrm{T}}\boldsymbol{x}_j \\ \text{subject to} & \boldsymbol{D}_j\boldsymbol{x}_j \le \sum_{p \in \mathcal{P}} \bar{\lambda}_p^j (\boldsymbol{D}_j \bar{\boldsymbol{x}}_j^p) + \sum_{r \in \mathcal{R}} \bar{\mu}_r^j (\boldsymbol{D}_j \tilde{\boldsymbol{x}}_j^r) \\ & \boldsymbol{A}_j\boldsymbol{x}_j \le \boldsymbol{b}_j \\ & \boldsymbol{x}_j \ge \boldsymbol{0} \qquad [X_j = \{\boldsymbol{x}_j \ge 0 \mid \boldsymbol{A}_j\boldsymbol{x}_j \le \boldsymbol{b}_j\}] \end{array}$$

since then $\sum_{j=1}^{n} \boldsymbol{D}_j\boldsymbol{x}_j \le \sum_{j=1}^{n} \boldsymbol{D}_j \underbrace{\left(\sum_{p \in \mathcal{P}} \bar{\lambda}_p^j \bar{\boldsymbol{x}}_j^p + \sum_{r \in \mathcal{R}} \bar{\mu}_r^j \tilde{\boldsymbol{x}}_j^r \right)}_{\boldsymbol{x}_j \in X_j} \le \boldsymbol{a}$

Estimates of the optimal objective value

$$\begin{aligned} z^* = \min \ & \sum_{p\in\mathcal{P}} \lambda_p(\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^p) + \sum_{r\in\mathcal{R}} \mu_r(\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}^r) \\ \text{s.t.} \ & \sum_{p\in\mathcal{P}} \lambda_p(\boldsymbol{A}\bar{\boldsymbol{x}}^p) + \sum_{r\in\mathcal{R}} \mu_r(\boldsymbol{D}\tilde{\boldsymbol{x}}^r) = \boldsymbol{d} \qquad | \ \boldsymbol{\pi} \\ & \sum_{p\in\mathcal{P}} \lambda_p = 1 \qquad | \ q \\ & \lambda_p, \mu_r \geq 0, \ p \in \mathcal{P}, r \in \mathcal{R} \end{aligned}$$

$$\begin{aligned} z^* \leq \bar{z} = \boldsymbol{b}^{\mathrm{T}}\bar{\boldsymbol{\pi}} + \bar{q} = \max_{(\boldsymbol{\pi},q)} \ & \boldsymbol{d}^{\mathrm{T}}\boldsymbol{\pi} + q \\ \text{s.t.} \ & (\boldsymbol{D}\bar{\boldsymbol{x}}^p)^{\mathrm{T}}\boldsymbol{\pi} + q \leq (\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^p), \quad p \in \bar{\mathcal{P}} \\ & (\boldsymbol{D}\tilde{\boldsymbol{x}}^r)^{\mathrm{T}}\boldsymbol{\pi} \leq (\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}^r), \quad r \in \bar{\mathcal{R}} \end{aligned}$$

Let λ_p^*, $p \in \mathcal{P}$, and μ_r^*, $r \in \mathcal{R}$, be optimal in the complete master problem, and $(\bar{\boldsymbol{\pi}}, \bar{q})$ an optimal dual solution for the columns in $\bar{\mathcal{P}}$ and $\bar{\mathcal{R}}$.

Multiply the right-hand side of the primal ($\boldsymbol{d}$ resp. 1) by $\bar{\boldsymbol{\pi}}$ resp. $\bar{q} \Longrightarrow$

$$\begin{aligned} & 0 \geq z^* - \bar{z} = z^* - \boldsymbol{b}^{\mathrm{T}}\bar{\boldsymbol{\pi}} - 1 \cdot \bar{q} = \sum_{p\in\mathcal{P}} \lambda_p^* \left[(\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^p) - (\boldsymbol{D}\bar{\boldsymbol{x}}^p)^{\mathrm{T}}\bar{\boldsymbol{\pi}} - \bar{q}\right] \\ & + \sum_{r\in\mathcal{R}} \mu_r^* \left[(\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}^r) - (\boldsymbol{D}\tilde{\boldsymbol{x}}^r)^{\mathrm{T}}\bar{\boldsymbol{\pi}}\right] \geq \min_{p\in\mathcal{P}} \left[(\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^p) - (\boldsymbol{D}\bar{\boldsymbol{x}}^p)^{\mathrm{T}}\bar{\boldsymbol{\pi}} - \bar{q}\right] \\ & + \sum_{r\in\mathcal{R}} \mu_r^* \min_{s\in\mathcal{R}} \left[(\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}^s) - (\boldsymbol{D}\tilde{\boldsymbol{x}}^s)^{\mathrm{T}}\bar{\boldsymbol{\pi}}\right] \end{aligned}$$

If the subproblem has an unbounded solution no optimistic estimate can be computed in this iteration; otherwise it holds that:

$$\min_{s\in\mathcal{R}} \left[(\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}^s) - (\boldsymbol{D}\tilde{\boldsymbol{x}}^s)^{\mathrm{T}}\bar{\boldsymbol{\pi}}\right] \geq 0$$

$\Longrightarrow$

$$\begin{aligned} \bar{z} \geq z^* \geq \bar{z} & + \min_{p\in\mathcal{P}} \left[(\boldsymbol{c} - \boldsymbol{D}^{\mathrm{T}}\bar{\boldsymbol{\pi}})^{\mathrm{T}}\bar{\boldsymbol{x}}^p - \bar{q}\right] \\ & = \bar{z} + \min_{\boldsymbol{x}\in X} (\boldsymbol{c} - \boldsymbol{D}^{\mathrm{T}}\bar{\boldsymbol{\pi}})^{\mathrm{T}}\boldsymbol{x} - \bar{q} \\ & = \underline{z} \end{aligned}$$

Convergence

The number of columns generated is finite, because X is polyhedral. When no more columns are generated, the solution to the last master problem will also solve the original linear problem. For each new column that is added to the master problem, its optimal objective value will decrease (or be kept constant). Hence, the pessimistic estimate $\bar{z}_k$ will converge monotonically to z^*.

The optimistic estimate $\underline{z}_k$ also converges, but perhaps not monotonically. If at iteration k an optimal solution to the complete master problem is received, $\underline{z}_k = \bar{z}_k$ holds.

Stopping criterion: $\bar{z}_k - \underline{z}_k^* \leq \varepsilon$, where $\underline{z}_k^* = \max_{s=1,\ldots,k} \underline{z}_s$ and $\varepsilon > 0$

A linear integer problem

$$z^* = \min \quad x_1 + 2x_2 \qquad x^* = (1,0),\ z^* = 1$$
$$\text{s.t.} \quad 2x_1 + 2x_2 \geq 1$$
$$x_1, x_2 \in \{0,1\},$$

$$z^*_{LP} = \min \quad x_1 + 2x_2 \qquad x^*_{LP} = \left(\frac{1}{2}, 0\right),\ z^*_{LP} = \frac{1}{2}$$
$$\text{s.t.} \quad 2x_1 + 2x_2 \geq 1$$
$$x_1, x_2 \in [0,1]$$

$z^*_{LP} \leq z^*$

Branch–and–price for linear 0/1 problems

$$\text{[IP]} \quad z^*_{\text{IP}} = \min\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$$
$$\text{s.t.}\ \boldsymbol{D}\boldsymbol{x} = \boldsymbol{d}$$
$$\boldsymbol{x} \in X = \{\boldsymbol{x} \in \mathbb{B}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\} = \{\bar{\boldsymbol{x}}^p \mid p \in \mathcal{P}\}$$

Inner representation (and convexification):

$$\text{conv}\, X = \left\{ \boldsymbol{x} = \sum_{p \in \mathcal{P}} \lambda_p \bar{\boldsymbol{x}}^p \,\middle|\, \sum_{p \in \mathcal{P}} \lambda_p = 1;\ \lambda_p \geq 0,\ p \in \mathcal{P} \right\}$$

Let $c_p = \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^p$ and $\boldsymbol{d}_p = \boldsymbol{D}\bar{\boldsymbol{x}}^p$, $p \in \mathcal{P}$.

Stronger formulation—Master problem

$$\text{[CP]} \quad z^*_{\text{IP}} = z^*_{\text{CP}} = \min \sum_{p \in \mathcal{P}} c_p \lambda_p$$
$$\text{s.t.} \sum_{p \in \mathcal{P}} \boldsymbol{d}_p \lambda_p = \boldsymbol{d}$$
$$\sum_{p \in \mathcal{P}} \lambda_p = 1$$
$$\lambda_p \in \{0,1\}, \qquad p \in \mathcal{P}$$

A continuous relaxation ($[\text{CP}^{cont}]$, to $\lambda_p \geq 0$) of [CP] gives the same lower bound as the Lagrangian dual for the constraints $\boldsymbol{D}\boldsymbol{x} = \boldsymbol{d}$. ($z^*_{LP} \leq z^{cont}_{CP} \leq z^*_{CP}$)

The continuous relaxation [LP] of [IP] is never better than any Lagrange dual bound.

Restricted master problem

Let $\bar{\mathcal{P}} \subseteq \mathcal{P}$

$$\overline{\text{CP}} \quad z^*_{\text{CP}} \geq z^{cont}_{\text{CP}} \leq \bar{z}_{\text{CP}} = \min \sum_{p \in \bar{\mathcal{P}}} c_p \lambda_p$$
$$\text{s.t.} \sum_{p \in \bar{\mathcal{P}}} \boldsymbol{d}_p \lambda_p = \boldsymbol{d}$$
$$\sum_{p \in \bar{\mathcal{P}}} \lambda_p = 1 \qquad (*)$$
$$\lambda_p \geq 0, \qquad p \in \bar{\mathcal{P}}$$

- Generate columns $\begin{pmatrix} c_p \\ \boldsymbol{d}_p \\ 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^p \\ \boldsymbol{D}\bar{\boldsymbol{x}}^p \\ 1 \end{pmatrix}$ until an (almost) optimal solution to [CP^{cont}], $\widehat{\lambda}_p$ ($p \in \bar{\mathcal{P}}$), is found
- $\widehat{\boldsymbol{x}} = \sum_{p \in \bar{\mathcal{P}}} \widehat{\lambda}_p \bar{\boldsymbol{x}}^p$

Branching over variable x_j with $0 < \hat{x}_j < 1$

$$x_j = 0 \qquad \text{or} \qquad x_j = 1$$

$$\Updownarrow \qquad\qquad \Updownarrow$$

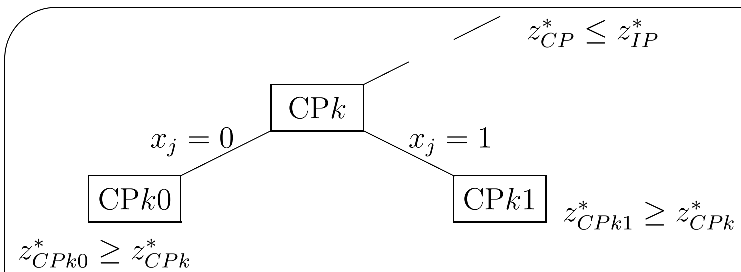
$$x_j = \sum_{p \in \bar{\mathcal{P}}} \lambda_p \bar{x}_j^p = 0 \qquad x_j = \sum_{p \in \bar{\mathcal{P}}} \lambda_p \bar{x}_j^p = 1$$

$$\Downarrow \qquad\qquad \Downarrow$$

$$\text{delete col's} \quad \sum_{p \in \bar{\mathcal{P}}: \bar{x}_j^p = 1} \lambda_p = 0 \qquad \sum_{p \in \bar{\mathcal{P}}: \bar{x}_j^p = 1} \lambda_p = 1 \quad \text{replaces } (*)$$

$$\Updownarrow \qquad\qquad \Updownarrow$$

$$\text{replaces } (*) \quad \sum_{p \in \bar{\mathcal{P}}: \bar{x}_j^p = 0} \lambda_p = 1 \qquad \sum_{p \in \bar{\mathcal{P}}: \bar{x}_j^p = 0} \lambda_p = 0 \quad \text{delete col's}$$



- In each node (CP, CP0, CP1, ...): Generate columns until (almost) optimal (all reduced costs ≥ 0) or verified infeasible
- If $\boldsymbol{x}^*_{CPk\ell\ldots}$ feasible $\Longrightarrow z^*_{CPk\ell\ldots} \geq z^*_{IP} \Longrightarrow$ Cut off the branch $(k, \ell, \ldots)$

$\Longrightarrow$ Cut branches $(r, s, \ldots)$ with $z^*_{CPrs\ldots} \geq z^*_{CPk\ell\ldots}$

The column generation subproblem, reduced costs

- $\min\limits_{\boldsymbol{x} \in X^k} (\boldsymbol{c} - \boldsymbol{D}^{\mathrm{T}}\widehat{\boldsymbol{\pi}}^k)^{\mathrm{T}}\boldsymbol{x} - \widehat{q}^k =: (\boldsymbol{c} - \boldsymbol{D}^{\mathrm{T}}\widehat{\boldsymbol{\pi}}^k)^{\mathrm{T}}\bar{\boldsymbol{x}}^p - \widehat{q}^k =: \bar{c}(\bar{\boldsymbol{x}}^p)$
- $(\widehat{\boldsymbol{\pi}}^k, \widehat{q}^k)$ is a dual solution to the RMP and $X^k = X \cap \{\boldsymbol{x} \mid x_j = k\}$, $k \in \{0, 1\}$ (etc. down the tree)
- If $\bar{c}(\bar{\boldsymbol{x}}^p) < 0$ then $\begin{pmatrix} \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}^p \\ \boldsymbol{D}\bar{\boldsymbol{x}}^p \\ 1 \end{pmatrix}$ is a new column in [CPk]
- Minimization? $\bar{\boldsymbol{x}}^r$ is good enough if $\bar{c}(\bar{\boldsymbol{x}}^r) < 0$
- If $\bar{c}(\bar{\boldsymbol{x}}^p) \geq 0$ then no more columns are needed to solve [CPk] to optimality.
- Same columns may be generated in different nodes $\Longrightarrow$ create "column pool" to check w.r.t. reduced costs $\bar{c}$

An instance solved by Branch–and–price

$$z^*_{IP} = \min \; x_1 + 2x_2 \quad \text{s.t. } 2x_1 + 2x_2 \geq 1, \; x_1, x_2 \in \{0,1\} \quad = z^*_{CP} \;\geq z^{cont}_{CP} = z^*_{LP} = \min \; x_1 + 2x_2 \quad \text{s.t. } 2x_1 + 2x_2 \geq 1, \; 0 \leq x_1, x_2 \leq 1$$

$$\text{conv}X = \text{conv}\left\{ \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix} \right\} = \left\{ \begin{pmatrix}\lambda_3+\lambda_4\\ \lambda_2+\lambda_4\end{pmatrix} \middle| \sum_{p=1}^{4} \lambda_p = 1; \; \lambda_p \geq 0 \; \forall p \right\}$$

[CP]

$$z^{cont}_{CP} = \min \; 2\lambda_2 + \lambda_3 + 3\lambda_4$$
$$\text{s.t. } 2\lambda_2 + 2\lambda_3 + 4\lambda_4 \geq 1$$
$$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$$
$$\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0$$

Start columns: λ_1 and λ_3

Choose e.g., $\begin{pmatrix}0\\0\end{pmatrix}$ and $\begin{pmatrix}1\\0\end{pmatrix}$, that is, the variables λ_1 and λ_3

$$z^{cont}_{CP} \leq \min \; \lambda_3 \quad \text{s.t. } 2\lambda_3 \geq 1, \; \lambda_1 + \lambda_3 = 1, \; \lambda_1, \lambda_3 \geq 0 \quad = \quad \max \; \pi + q \quad \text{s.t. } q \leq 0, \; 2\pi + q \leq 1, \; \pi \geq 0$$

Solution: $(\widehat{\lambda}_1, \widehat{\lambda}_3) = (\frac{1}{2}, \frac{1}{2}) \Longrightarrow \widehat{\boldsymbol{x}} = (\frac{1}{2}, 0)^{\mathrm{T}}$, $\widehat{\pi} = \frac{1}{2}$, $\widehat{q} = 0$
Reduced costs: $\min_{\boldsymbol{x} \in [0,1]^2} \{(0,1)\boldsymbol{x}\} = 0 \Longrightarrow$ Optimum for CP!

Fixations: $x_1 = 0$ or $x_1 = 1$

$\Downarrow$ $\quad\Downarrow$

$\lambda_3 = 0 \quad \lambda_1 = 0$

Branching, left (CP0): $\lambda_3 = 0$

$$\min \; 0 \quad \text{s.t. } 0 \geq 1, \; \lambda_1 = 1, \; \lambda_1 \geq 0 \Longrightarrow \begin{bmatrix}\text{infeasible}\\ \Downarrow \\ \text{add} \\ \text{column}\end{bmatrix} \Longrightarrow z_{CP0} \leq \min \; 2\lambda_2 \quad \text{s.t. } 2\lambda_2 \geq 1, \; \lambda_1 + \lambda_2 = 1, \; \lambda_1, \lambda_2 \geq 0$$

$$= \max \; \pi + q \quad \text{s.t. } q \leq 0, \; 2\pi + q \leq 2, \; \pi \geq 0$$

Solution: $(\widehat{\lambda}_1, \widehat{\lambda}_2) = (\frac{1}{2}, \frac{1}{2})$
$\Longrightarrow \widehat{\boldsymbol{x}} = (0, \frac{1}{2})^{\mathrm{T}}$
$\widehat{\pi} = 1, \quad \widehat{q} = 0$

Reduced costs: $\min_{\boldsymbol{x} \in [0,1]^2} \{(-1,0)\boldsymbol{x} - 0\} = -1 < 0$
$\Longrightarrow$ New column! (λ_3 or λ_4, but $\lambda_3 \equiv 0$) $\Longrightarrow$ Choose λ_4

$$z_{CP0} \leq \min \; 2\lambda_2 + 3\lambda_4 \quad \text{s.t. } 2\lambda_2 + 4\lambda_4 \geq 1, \; \lambda_1 + \lambda_2 + \lambda_4 = 1, \; \lambda_1, \lambda_2, \lambda_4 \geq 0 \quad = \max \; \pi + q \quad \text{s.t. } q \leq 0, \; 2\pi + q \leq 2, \; 4\pi + q \leq 3, \; \pi \geq 0$$

- Solution: $(\widehat{\lambda}_1, \widehat{\lambda}_3, \widehat{\lambda}_4) = (\frac{3}{4}, 0, \frac{1}{4}) \Longrightarrow \widehat{\boldsymbol{x}} = (\frac{1}{4}, \frac{1}{4})^{\mathrm{T}}$, $\widehat{\pi} = \frac{3}{4}$, $\widehat{q} = 0$
- Reduced costs: $\min_{\boldsymbol{x} \in [0,1]^2} \{(-\frac{1}{2}, \frac{1}{2})\boldsymbol{x}\} = -\frac{1}{2} \Longrightarrow$
- Generate new column: λ_3, but $\lambda_3 \equiv 0 \Longrightarrow$ Optimum for CP0

Branching, right (CP1): $\lambda_1 = 0$

$$z_{CP1} \le \min \ \lambda_3 \quad \text{s.t.} \ 2\lambda_3 \ge 1, \ \lambda_3 = 1, \ \lambda_3 \ge 0 \quad = \quad \max \ \pi + q \quad \text{s.t.} \ 2\pi + q \le 1, \ \pi \ge 0$$

- Solution: $\widehat{\lambda}_3 = 1 \Longrightarrow \widehat{\boldsymbol{x}} = (1,0)^{\mathrm{T}}$, $\widehat{\pi} = 0$, $\widehat{q} = 1$
- Reduced costs: $\min_{\boldsymbol{x}\in[0,1]^2} \{(1,2)\boldsymbol{x} - 1\} = -1 < 0 \Longrightarrow$
- Generate new column: λ_1, but $\lambda_1 \equiv 0 \Longrightarrow$ Optimum for CP1 !!

Branching, left, left: (CP00) $\lambda_2 = \lambda_4 = 0$

CP00: $\lambda_2 = \lambda_3 = \lambda_4 = 0 \Longrightarrow$ infeasible

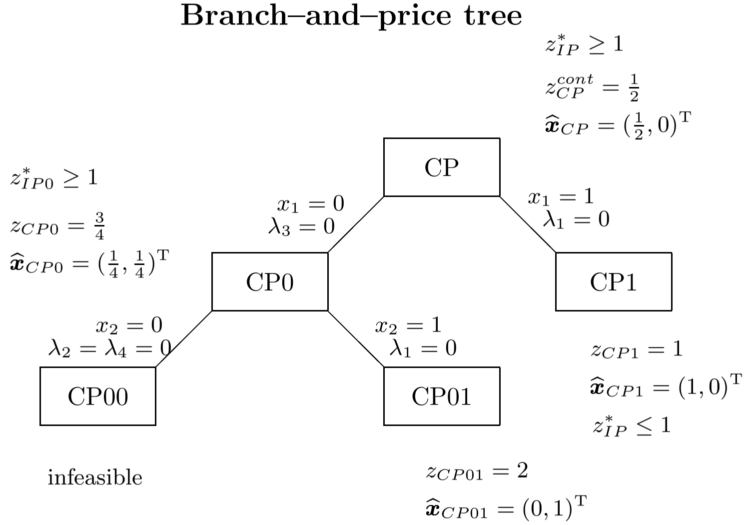
Branching, left, right: (CP01) $\lambda_1 = 0$

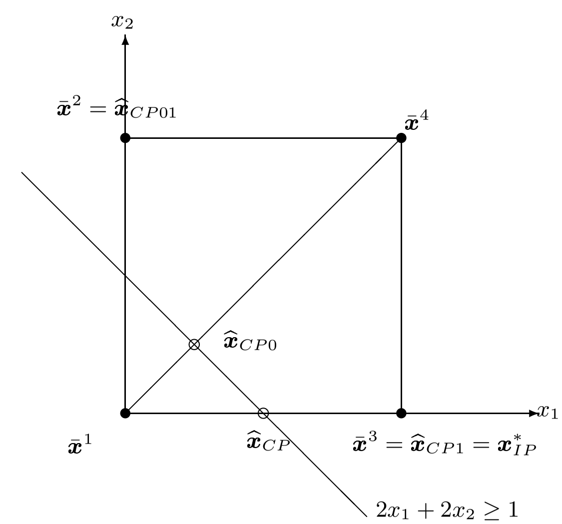
CP01: $\lambda_1 = \lambda_3 = 0$

$$z_{CP01} \le \min \ 2\lambda_2 + 3\lambda_4 \quad \text{s.t.} \ 2\lambda_2 + 4\lambda_4 \ge 1, \ \lambda_2 + \lambda_4 = 1, \ \lambda_2, \lambda_4 \ge 0 \quad = \quad \max \ \pi + q \quad \text{s.t.} \ 2\pi + q \le 2, \ 4\pi + q \le 3, \ \pi \ge 0$$

- Solution: $(\widehat{\lambda}_2, \widehat{\lambda}_4) = (1,0)^{\mathrm{T}} \Longrightarrow \widehat{\boldsymbol{x}} = (0,1)^{\mathrm{T}}$, $\widehat{\pi} = 0$, $\widehat{q} = 2$
- Reduced costs: $\min_{\boldsymbol{x}\in[0,1]^2} \{(1,2)\boldsymbol{x} - 2\} = -2 < 0$
  $\Longrightarrow$ Generate new column: λ_1, but $\lambda_1 \equiv 0$
  $\Longrightarrow$ Generate new column: λ_3, but $\lambda_3 \equiv 0$
  $\Longrightarrow$ Optimum for CP01 !!

Branch–and–price tree





Benders decomposition for mixed-integer linear problems—Lasdon (1970)

- Model:

$$\begin{aligned} \text{minimum} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + f(\boldsymbol{y}), \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{F}(\boldsymbol{y}) \geq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}^n, \quad \boldsymbol{y} \in S. \end{aligned}$$

- The variables $\boldsymbol{y}$ are “difficult” because
  - the set S may be complicated, like $S \subseteq \{0,1\}^p$;
  - f and/or $\boldsymbol{F}$ may be nonlinear;
  - the vector $\boldsymbol{F}(\boldsymbol{y})$ may cover every row, while the problem in $\boldsymbol{x}$ for fixed $\boldsymbol{y}$ may separate;
  - the problem in $\boldsymbol{x}$ is *linear*.

- Typical application: Multi-stage stochastic programming. Choose $\boldsymbol{y}$ such that an expected cost over time is minimized; uncertainty in data is translated into future scenarios and variables $\boldsymbol{x}$ representing future activities that “adjust” the $\boldsymbol{y}$ that was chosen before knowledge of the values of the stochastic variables has been revealed. The $\boldsymbol{y}$ should therefore be chosen such that the expected value of the future optimization over $\boldsymbol{x}$ is the best.

- Idea: Temporarily fix $\boldsymbol{y}$, solve the remaining problem over $\boldsymbol{x}$ parameterized over $\boldsymbol{y}$. Utilize the structure of the problem to improve the guess of an optimal value of $\boldsymbol{y}$. Repeat.
- Similar to solving the problem of minimizing a function η over two vectors $(\boldsymbol{v}, \boldsymbol{w})$ as follows:

$$\inf_{(\boldsymbol{v},\boldsymbol{w})} \eta(\boldsymbol{v},\boldsymbol{w}) = \inf_{\boldsymbol{v}} \xi(\boldsymbol{v}), \text{ where } \xi(\boldsymbol{v}) = \inf_{\boldsymbol{w}} \eta(\boldsymbol{v},\boldsymbol{w}), \ \boldsymbol{v} \in \mathbb{R}^m.$$

- In effect, we substitute the variable $\boldsymbol{w}$ by always minimizing over it, and work with the remaining problem in $\boldsymbol{v}$.

- Benders decomposition centers on the possibility to construct an approximation of this problem over $\boldsymbol{v}$ by utilizing LP duality.
- In the case that the problem over $\boldsymbol{y}$ also is linear we recover the cutting plane methods from above. Benders decomposition is more general however, because we can solve problem that have a positive duality gap. In other words, the workings of Benders decomposition does not rely on the existence of optimal Lagrange multipliers and strong duality.

The Benders sub- and master problems

- Which $\boldsymbol{y}$ are feasible? We must choose $\boldsymbol{y} \in S$ such that the remaining problem in $\boldsymbol{x}$ is feasible. In other words: choose $\boldsymbol{y}$ in the set

$$R := \{ \boldsymbol{y} \in S \mid \exists \boldsymbol{x} \geq \boldsymbol{0}^n \text{ with } \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y}) \}$$

We apply Farkas' Lemma to this system, or rather to the equivalent system (with $\boldsymbol{y}$ fixed)

$$\boldsymbol{A}\boldsymbol{x} - \boldsymbol{s} = \boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y}), \quad (4a)$$
$$\boldsymbol{x} \geq \boldsymbol{0}^n, \quad (4b)$$
$$\boldsymbol{s} \geq \boldsymbol{0}^m. \quad (4c)$$

- From Farkas' Lemma, $\boldsymbol{y} \in R$ is and only if

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{u} \leq \boldsymbol{0}^n, \ \boldsymbol{u} \geq \boldsymbol{0}^m \quad \Longrightarrow \quad [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u} \leq 0,$$

in other words,

$$[\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u}_i^r \leq 0$$

holds for every extreme ray $\boldsymbol{u}_i^r$, $i = 1, \ldots, n_r$ of the polyhedral cone $C = \{ \boldsymbol{u} \in \mathbb{R}_+^m \mid \boldsymbol{A}^{\mathrm{T}}\boldsymbol{u} \leq \boldsymbol{0}^n \}$.

- We here made good use of the Representation Theorem for a polyhedral cone.
- Given $\boldsymbol{y} \in R$, the optimal value in Benders' subproblem

is

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimum}} \ & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to} \ & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y}), \\ & \boldsymbol{x} \geq \boldsymbol{0}^n, \end{aligned}$$

which by LP duality equals

$$\begin{aligned} \underset{\boldsymbol{u}}{\text{maximum}} \ & [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u}, \\ \text{subject to} \ & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{u} \leq \boldsymbol{c}, \\ & \boldsymbol{u} \geq \boldsymbol{0}^m, \end{aligned}$$

provided that the first problem does not have an infinite solution.

- We prefer the dual formulation, since its constraints do not depend on $\boldsymbol{y}$; moreover, the extreme rays of its feasible set are given by the vectors $\boldsymbol{u}_i^r$, $i = 1, \ldots, n_r$, discussed above. Let $\boldsymbol{u}_i^p$, $i = 1, \ldots, n_p$, denote the extreme points of this set.
- This completes the subproblem. Let's now study the restricted master problem of Benders' algorithm.

- The original problem is equivalent to the problem to

$$\min_{\boldsymbol{y}\in R} \left\{ f(\boldsymbol{y}) + \max_{\boldsymbol{u}} \{ [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u} \mid \boldsymbol{A}^{\mathrm{T}}\boldsymbol{u} \leq \boldsymbol{c};\ \boldsymbol{u} \geq \boldsymbol{0}^m \} \right\}$$

$$= \min_{\boldsymbol{y}\in R} \left\{ f(\boldsymbol{y}) + \max_{i=1,\dots n_p} \{ [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u}_i^p \} \right\}$$

$$\begin{aligned} = \min\ & z \\ \text{s.t. } & z \geq f(\boldsymbol{y}) + [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u}_i^p, \quad i = 1,\dots,n_p, \\ & \boldsymbol{y} \in R, \end{aligned}$$

$$\begin{aligned} = \min\ & z \\ \text{s.t. } & z \geq f(\boldsymbol{y}) + [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u}_i^p, \quad i = 1,\dots,n_p, \\ & 0 \geq [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u}_i^r, \quad i = 1,\dots,n_r, \\ & \boldsymbol{y} \in S. \end{aligned}$$

- Suppose then that not the whole sets of constraints in the latter problem is known, and replace "$i = 1,\dots,n_p$" with "$i \in I_1$", respectively "$i = 1,\dots,n_r$" with "$i \in I_2$," where $I_1 \subset \{1,\dots,n_p\}$ and $I_2 \subset \{1,\dots,n_r\}$.
- Since not all constraints are included, we get a lower bound on the optimal value of the original problem. Suppose then that $(z^0, \boldsymbol{y}^0)$ is a finite optimal solution to this problem. In order to check if this is indeed an optimal solution to the original problem, we check for the most violated constraint, which we either satisfy (thus having established that $\boldsymbol{y}^0$ indeed is optimal) or, if not, we include this new constraint, improving either the set I_1 or I_2, and possibly improving the lower

bound.

- The search for a new constraint is of course the same as solving the dual of Benders' subproblem with $\boldsymbol{y} = \boldsymbol{y}^0$!
- This problem gives us a feasible solution to the original problem, and therefore also an upper bound, provided that it is finite.
- If this problem has an unbounded solution, then it is unbounded along an extreme ray: $[\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y}^0)]^{\mathrm{T}}\boldsymbol{u}_i^r > 0$. We then add the constraint $0 \geq [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u}_i^r$ to the RMP (enriching the set I_2).

- Suppose instead that we find a finite optimal solution. Let $\boldsymbol{u}_i^p$ be an optimal extreme point. If it holds that $z^0 < f(\boldsymbol{y}^0) + [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y}^0)]^{\mathrm{T}}\boldsymbol{u}_i^p$, we add the constraint $z \geq [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y})]^{\mathrm{T}}\boldsymbol{u}_i^p$ to the description of the RMP (enriching I_1).
- If however $z^0 \geq f(\boldsymbol{y}^0) + [\boldsymbol{b} - \boldsymbol{F}(\boldsymbol{y}^0)]^{\mathrm{T}}\boldsymbol{u}_i^p$ then in fact equality holds in this inequality ($>$ can never happen—why?). We have then identified an optimal solution to the original problem, and terminate.

Convergence

- Suppose that S is closed and bounded and that f and $\boldsymbol{F}$ both are continuous on S. Then provided that the computations are exact we terminate in a finite number of iterations with an optimal solution.
- Proof is by the finiteness of the number of constraints in the complete master problem, that is, the number of extreme points and rays in any polyhedron.
- A numerical example of the use of Benders decomposition is found in Lasdon (1970, Sections 7.3.3–7.3.5).

- Note the resemblance to the Dantzig–Wolfe algorithm! In fact, if f and $\boldsymbol{F}$ both are linear, then they coincide, in the sense that their subproblems and restricted master problems are identical!
- Modern implementations of the Dantzig–Wolfe and Benders algorithms are inexact, that is, at least their RMP:s are not solved exactly. Moreover, their RMP:s are often restricted such that there is an additional "box constraint" added. This constraint forces the solution to the next RMP to be relatively close to the previous one. The effect is that of a stabilization; otherwise, there is a risk that the sequence of solutions to the RMP:s "jump about," and convergence becomes

slow as the optimal solution is approached. This was observed quite early on with the Dantzig–Wolfe algorithm, which even can be enriched with non-linear "penalty" terms in the RMP to further stabilize convergence. In any case, convergence holds also under these modifications, except perhaps for the finiteness.