5. Column generation

Lec 9, 10.1.13

Notiztitel 11.01.2013

rough idea: suppose we have an LP with a huge number of variables, let's ignore some of them and add them later.

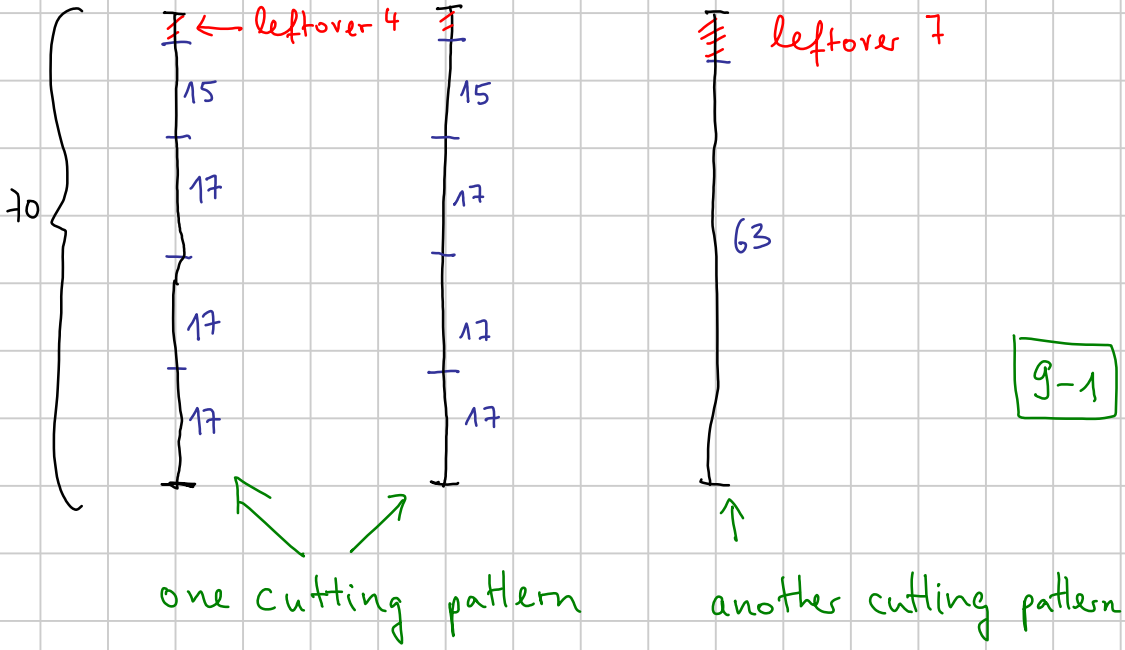
Example 5.1 Cutting stock problem (CSP)

given:
- supply of long tubes, (or rolls of paper) each of length $W \in \mathbb{N}$
- demand: $d_i \in \mathbb{N}$ short tubes of length $w_i \leq W$ for $i \in [n]$
- where short tubes must be cut out of long tubes

minimize:
- number of long tubes needed

e.g: $W = 70$, $w_1 = 17$ $d_1 = 6$, $w_2 = 15$ $d_2 = 2$, $w_3 = 63$ $d_3 = 1$

a feasible solution:



A cutting pattern is a vector $x \in \mathbb{N}_0^n$, where x_i denotes the number of short tubes of length w_i cut out of a long tube.

feasibility: $\sum_{i=1}^{n} x_i w_i \leq W$

number of long tubes needed $\leq \sum_{i=1}^{n} d_i =: K$

define $y \in \{0,1\}^K$ where for $k \in [K]$

$y_k = 1$ iff k-th long tube is needed

$\rightsquigarrow$ goal: minimize $\sum_{k=1}^{K} y_k$

For the k-th long tube denote by $x^k \in \mathbb{N}_0^n$ the (unknown) cutting pattern

$\rightsquigarrow$ feasibility: $\sum_{i=1}^{n} x_i^k w_i \leq W \cdot y_k \quad \forall k \in [K]$

$\rightsquigarrow$ demands: $\sum_{k=1}^{K} x_i^k \geq d_i \quad \forall i \in [n]$

$$ILP_1: \left\{ \min_{y,x} \sum_{k=1}^{K} y_k : y \in \{0,1\}^K \right.$$

$$x^k \in \mathbb{N}_0^n \quad \text{for } k = 1, \dots, K$$

$$\sum_{k=1}^{K} x_i^k \geq d_i \quad \text{for } i = 1, \dots, n$$

$$\left. \sum w_i x_i^k \leq W y_k \quad \text{for } k = 1, \dots, K \right\}$$

advantage: small number of variables and constraints

but: gap between ILP_1 and LP-relaxation can be big.

Prop. 5.2 Consider the LP-relaxation LP_1 of ILP_1, where $x_i^k, y_k \in \mathbb{R}$.

Then its optimal value is $\frac{\sum_{i=1}^{n} w_i d_i}{W}$.
(number of long tubes)

Proof: obvious.

Example 5.3 $n := 1$

d_1 arbitrary

$w_1 := \left\lfloor \frac{W}{2} \right\rfloor + 1$

Then optimal value of ILP_1 is d_1, but by Prop 5.2 the optimal value of LP_1 is

$$\frac{d_1 w_1}{W} = \frac{d_1 \left(\left\lfloor \frac{W}{2} \right\rfloor + 1 \right)}{W} \approx \frac{d_1}{2}.$$

$\rightsquigarrow$ big gap.

Example 5.4 We try a second ILP-formulation

$Q :=$ set of all valid cutting patterns

$Q = \{ x^j : j = 1, \dots, |Q| \}$. (This is a huge set.)

Let $\mu_j \in \mathbb{N}_0$ denote how often pattern $x^j \in Q$ is used.

$$ILP_2: \quad \Big\{ \min \sum_{j=1}^{|Q|} \mu_j \;:\; \mu_j \in \mathbb{N}_0 \quad \forall j=1,\dots,|Q| \;;\; \sum_{j=1}^{|Q|} x_i^j \mu_j \geq d_i \quad \forall i=1,\dots,n \Big\}$$

(Master problem)

Here we think of the set $Q = \{x^j : j \in [|Q|]\}$ as given. We know that

(subproblem) $$Q = \Big\{ x \in \mathbb{N}_0^n : \sum_{i=1}^{n} w_i x_i \leq W \Big\}.$$

Column generation approach to solve ILP_2:

Start with the master problem over a small subset of Q and solve it. If this solution is not optimal for the original problem (how do we check this?), generate additional cutting patterns from Q (which?).

To answer the two blue questions, we first recall basic facts about the simplex method.

LP: $$\max \{ c^T x : Ax = b,\ x \geq 0 \}$$
$$A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m,\ c \in \mathbb{R}^n$$

$$\left.\begin{aligned} B &= (p_1,\dots,p_m) \in [n]^m \\ N &= (q_1,\dots,q_{n-m}) \in [n]^{n-m} \end{aligned}\right\} \text{ vectors with column indices}$$
such that $\{p_1,\dots,p_m\} \dot\cup \{q_1,\dots,q_{n-m}\} = [n]$.

Denote by A_B and A_N the submatrices of A consisting of the columns of A, indexed by the indices in B and N.

If A_B regular, then (A_B, A_N) is called a basis
x defined by $x = (x_B, x_N)$ with

$$x_B := A_B^{-1} b \quad \text{and} \quad x_N := 0$$

is called the corresponding basis solution.

Then

$$c^T x = c_B^T A_B^{-1} b + \underbrace{\left(c_N^T - c_B^T A_B^{-1} A_N\right)}_{\text{reduced costs: } \bar{c}^T} x_N$$

If $\bar{c} \leq 0$, then the basis solution is optimal.
If $\bar{c}_j > 0$, then the j-th column of A_N should enter the basis.

Column generation and Simplex method:

LP: $\max \{ c^T x : Ax \leq b, \; x \geq 0 \}$

Let A' consist of, say, the first n' columns of A, shrink x to x' and c to c' accordingly.

LP': $\max \{ c'^T x' : A' x' \leq b, \; x' \geq 0 \}$

Let $x^{*\prime}$ be an optimal solution of LP' with basis (A'_B, A'_N).
Then $(x^{*\prime}_B, 0)$ is feasible for LP and defines a feasible basis (A'_B, A_N) which may or may not be optimal for LP.

Check this by looking at the reduced cost

$$\bar{c}^T := c_N^T - c_B^T A_B'^{-1} A_N$$

A column in A_N with positive reduced costs is a candidate to enter the basis for LP.

We find the best such candidate by taking the one with largest reduced costs by solving the

pricing problem: $\max \{ c_j - y_B^T a^j : a^j$ is a column of A_N, j the corresponding index $\}$

where $y_B^T := c_B^T A_B'^{-1}$.

Suppose we can solve the pricing problem efficiently (i.e. without looking at all the columns).

Then this gives us the column with the largest reduced costs, add it to A' and iterate. Stop when $\bar{c} \leq 0$, because we have then found an optimal solution.

What does the pricing problem look like in our ILP_2-case?

Let A_B' be an optimal basis of the current (related) master problem.

9-6

Then a column of A_N is just a

feasible cutting pattern, ie. a vector $x \in \mathbb{N}_0^n$ satisfying the constraints from the subproblem:

$$\sum_{i=1}^{n} w_i x_i \leq W .$$

Hence the pricing problem becomes

$$\min \left\{ 1 - y_B^T x : \sum_{i=1}^{n} w_i x_i \leq W, \; x \in \mathbb{N}_0^n \right\}$$

↑ because CSP is a min!-problem

where $y_B^T := c_B^T A_B^{-1}$

and this is (nothing but a good old) Knapsack problem.

Although Knapsack is NP-hard, it is usually well-behaved and can be solved efficiently in practice (dyn. progr., approximation; see Lectures 2 and 3).
Also note that $n \ll |Q|$,
so we have reduced the complexity a lot.