SOLVING A NETWORK DESIGN PROBLEM USING BENDERS DECOMPOSITION

Benders Decomposition may be applied to linear and mixed integer linear problems with this structure:

$$
\begin{array}{rl}
Min_{\,x,\,y} & c^\top x + f^\top y \\
s.t: & Dx + Fy = d \\
 & x \geq 0, \\
\texttt{(P)} & y \in Y
\end{array}
$$

where variables x are may take continuous values and variables y (the linking variables) may be continuous or discrete (Y may be a polyhedron $P \subset R^p$ or $Y = P \cap Z^p$).

Assume that variables y are fixed to a given value $\bar{y}$. Then problem (P) reduces to a linear programming problem:

$$
\begin{array}{rl}
Min_{\,x} & z_P = c^\top x \\
s.t: & Dx = d - F\bar{y} \\
\texttt{(PRIMAL)} & x \geq 0
\end{array}
\quad \Leftrightarrow \quad
\begin{array}{rl}
Max_{\,u} & z_D = (d - F\bar{y})^\top u \\
\texttt{(DUAL)} & D^\top u \leq c
\end{array}
$$

If a collection of vertexes $\hat{v}^i$, $i = 1, 2, ...\mu$ and extreme rays $\overrightarrow{w}^j$, $j = 1, 2, ...\nu$ of the polyhedron defining the feasible region of the dual problem (DUAL) has been generated, then the initial problem (P) can be approximated by the following problem:

$$
\begin{array}{rll}
Min_{\,z,\,y} & z & \\
s.t: & f^\top y + (d - Fy)^\top \hat{v}^i - z \leq 0, & i = 1, 2, ..., \mu \\
 & (d - Fy)^\top \overrightarrow{w}^j \leq 0, & j = 1, 2, ..., \nu \\
\texttt{(MASTER PROBLEM)} & y \in Y &
\end{array}
$$

After being solved, the *master problem* provides an approximation to the optimal value of the objective function of the original problem (P), $\bar{z}$, and a value $\bar{y}$ for the link variables. This value $\bar{y}$ for the linking variables can be used to solve the dual problem (DUAL) referred to as the *subproblem*, which will generate new vertexes and rays $\hat{v}, \overrightarrow{w}$, that will result in new constraints for the master problem and so on.

At iteration ℓ:

1. Solve the master problem $\rightarrow \bar{z}, \bar{y}$.
2. Solve the subproblem:
   (a) The subproblem DUAL has no solution because $D^\top u \leq c$ is an empty polyhedron; In this case STOP because problem (P) has no solution or it is unbounded.
   (b) Subproblem DUAL has a solution;
      i. $z_D = \infty$. Then, if the simplex method has been used, an extreme ray $\overrightarrow{w}^*$ has been obtained so that $(d - Fy)^\top \overrightarrow{w}^* > 0$.
      Thus, constraint $(d - Fy)^\top \overrightarrow{w}^* \leq 0$ needs to be added to the *master problem*, which will be solved again in a new iteration.
      ii. $z_D < \infty$. A vertex $\hat{v}^*$ will be generated after solving problem DUAL using the simplex method.
         A.) If $f^\top \bar{y} + (d - F\bar{y})^\top \hat{v}^* - \bar{z} \leq 0$. then STOP.
         (In this case: $\bar{z} \geq f^\top \bar{y} + (d - F\bar{y})^\top \hat{v}^* \geq f^\top \bar{y} + (d - F\bar{y})^\top \hat{v}^\ell$, $1 \leq \ell \leq \mu$. Also, $(d - F\bar{y})^\top w^p \leq 0$, $1 \leq p \leq \nu$.)
         Optimal solution of problem: objective function $= \bar{z}$, linking variables $= \bar{y}$, dual variables $= \hat{v}^*$.

         As a practical stopping criterion adopt: $f^\top \bar{y} + (d - F\bar{y})^\top \hat{v}^* - \bar{z} \leq \epsilon$ or, better, using a relative gap: $f^\top \bar{y} + (d - F\bar{y})^\top \hat{v}^* - \bar{z} \leq \varepsilon \bar{z}$

B.) $f^\top \bar{y} + (d - F\bar{y})^\top \hat{v}^* - \bar{z} > 0$.
Then, constraint $f^\top y + (d - Fy)^\top \hat{v}^* - z \leq 0$ needs to be added to the *master problem*, which will be solved again in a new iteration.

The previous scheme extends readily to the case in which variables y are integer. Then, the *master problem* becomes a mixed linear-integer programming problem with only one continuous variable.

Network Design. Description of the Assignment

Consider the problem consisting in deciding which links in a transportation or in a data communications network are necessary by balancing investment costs and reductions in exploitation costs. Assume that initially the network presentes a configuration given by $G' = (N, A')$ and that there exists the possibility of increasing the number of network links taking the candidate links from a pre-specified set $\hat{A}$, so that the final network may become $G = (N, A)$ with $A = A' \cup \hat{A}$. The question is then which links in $\hat{A}$ are suitable to be added so that the total cost (investment+exploitation) is the minimum possible. The cost will be made up by two components:

a) on the one hand, adding each new link $a \in \hat{A}$ presents a fix cost f_a (purchase+instalation).

b) when in service, a link $a \in \hat{A}$ will have a cost per each unit of flow crossing that link during a given amortization period. Assume is addition, that these cost per unit depend on the starting origin (origin) of the trip.

In this way, if K denotes the set of origins in the network, the cost throughout the amortization period can be formulated as:

$$\text{total cost} = \sum_{\ell \in K} \sum_{a \in A} c_a^\ell x_a^\ell + \sum_{a \in \hat{A}} f_a y_a$$

being $y_a = 1$ if link $a \in \hat{A}$ is added to the network and 0 otherwise.

The problem will be now formulated using matrix notation.

Let B the node-link incidence matrix of the network, x^ℓ the vector of flows originating at $\ell \in K$. Assume that g_ℓ is the total flow outgoing from origin $\ell \in K$ and that if $D(\ell)$ is the set of destinations corresponding to origin ℓ, then $g_{\ell,j}$ is the flow arriving at $j \in D(\ell)$, in such a way that:

$$g_\ell = \sum_{j \in D(\ell)} g_{\ell,j}$$

Let vector t^ℓ have so many components as nodes in the network. Component t_i^ℓ corresponding to node $i \in N$ is:
$t_i^\ell = -g_{\ell,i}$ if $i \neq \ell$ i $t_i^\ell = g_\ell$ if node i is precisely the origin ℓ.
Then x^ℓ obey to the following relationships:

$$Bx^\ell = t^\ell, \; x^\ell \geq 0$$

(See a complete example in the ending page).
The problem of network design can be formulated :

$$(1) \qquad \begin{array}{ll} Min_{\,x,y} & \sum_{\ell \in K} {c^\ell}^\top x^\ell + f^\top y \\ (A) & Bx^\ell = t^\ell, \;\; \ell \in K \\ (B) & x_a^\ell \leq \rho\, y_a, \;\; \ell \in K, \; a \in \hat{A} \\ (C) & x^\ell \geq 0 \\ & y \in \{0,1\}^{|\hat{A}|} \end{array}$$

and in scalar notation:

$$\begin{array}{lll} Min_{\,x,y} & \sum_{\ell \in K} \sum_{a \in A} c_a^\ell x_a^\ell + \sum_{a \in \hat{A}} f_a y_a & \\ & \sum_{r \in I(i)} x_{r,i}^\ell \; - \; \sum_{s \in E(i)} x_{i,s}^\ell \; = \; t_i^\ell, & i \in N, \; \ell \in K \\ & x_a^\ell \leq \rho\, y_a, & a \in \hat{A}, \; \ell \in K \\ & x_a^\ell \geq 0, & a \in A, \; \ell \in K \\ & y_a \in \{0,1\}, & a \in \hat{A} \end{array}$$

Constant ρ must be chosen big enough so that if $y_a = 1$ then the maximum value that can be achieved by x_a^ℓ be smaller to ρ. Adopt $\rho > \sum_{\ell \in K} g_\ell$.

In order to write the *master problem mestre*, constraints (A), (B) can be stated in matrix notation as (A'), (B'):

$$(A'),\ (B') \quad \left(\begin{array}{cccc|c} \mathcal{B} & 0 & \dots & 0 & -F \\ 0 & \mathcal{B} & \dots & 0 & -F \\ \vdots & \vdots & \ddots & \vdots & -F \\ 0 & 0 & \dots & \mathcal{B} & -F \end{array}\right) \left(\begin{array}{c} \mathrm{x}^{\ell_1} \\ \mathrm{x}^{\ell_2} \\ \vdots \\ \mathrm{x}^{\ell_{|K|}} \\ \hline y \end{array}\right) = \left(\begin{array}{c} \mathcal{T}^{\ell_1} \\ \mathcal{T}^{\ell_2} \\ \vdots \\ \mathcal{T}^{\ell_{|K|}} \end{array}\right), \quad \mathrm{x}^\ell \geq 0,\ \ell \in K$$

where:

$$\mathcal{B} = \left(\begin{array}{c|c} B & 0 \\ \hline I_{\hat{A}}|0 & I_{\hat{A}} \end{array}\right), \quad \mathrm{x}^\ell = \left(\begin{array}{c} x^\ell \\ \hline \sigma^\ell \end{array}\right), \mathcal{T}^\ell = \left(\begin{array}{c} t^\ell \\ \hline 0 \end{array}\right), \quad F = \rho \left(\begin{array}{c} 0 \\ \hline I_{\hat{A}} \end{array}\right)$$

In previous expressions $I_{\hat{A}}$ denotes the identity matrix of size $|\hat{A}|$.

Constraints (A') will be rewritten by blocks as well the corresponding dual variables:

$$\mathcal{B}\mathrm{x}^\ell - Fy = \mathcal{T}^\ell \ \mid \ \theta^\ell$$

equivalently:

(2)
$$\left(\begin{array}{c|c} B & 0 \\ \hline I_{\hat{A}}|0 & I_{\hat{A}} \end{array}\right) \left(\begin{array}{c} x^\ell \\ \hline \sigma^\ell \end{array}\right) - \rho \left(\begin{array}{c} 0 \\ \hline I_{\hat{A}} \end{array}\right) = \left(\begin{array}{c} t^\ell \\ \hline 0 \end{array}\right) \ \Bigg| \ \left(\begin{array}{c} u^\ell \\ \hline \tau^\ell \end{array}\right) (= \theta^\ell)\ ,\ \ell \in K$$

In scalar notation:

$$\begin{array}{lll} Bx^\ell & = t^\ell & |u^\ell \\ x_a^\ell + \sigma_a^\ell - \rho y_a & = 0 & |\tau_a^\ell \end{array}$$

Thus, it must be taken into account that those variables intervening in formulating the Benders cuts are $\theta^\ell = (u^\ell, \tau^\ell)$. These dual variables are the ones associated to constraints (A'), (B'). Instead, it is convenient to solve the subproblem in the following primal form, which decomposes in $|K|$ subproblems.

$$\begin{array}{lll} Min_{\ x,y} & c^{\ell\top} x^\ell & \\ & Bx^\ell = t^\ell, \quad |w^\ell & \\ (\mathtt{SUBP}\ \ \ell) & x_a^\ell \leq \rho\, \bar{y}_a, \quad |\tau_a^{\ell -}, \ \ a \in \hat{A} & \rightarrow \ (\hat{x}^\ell, \hat{u}^\ell, \hat{\tau}^{\ell +}, \hat{\tau}^{\ell -}) \\ & x^\ell \geq 0, \quad\quad |\tau^{\ell +} & \end{array}$$

It is necessary to take parameter $\rho > \sum_{\ell \in K} g_\ell$. The relationship between multipliers $w^\ell, \tau_a^{\ell -}, \tau^{\ell +}$ for problem (SUBP ℓ) and $\theta^\ell = (u^\ell, \tau^\ell)$ is the following one:

(3)
$$w_i^\ell = u_i^\ell, \qquad i \in N$$

(4)
$$\tau_a^\ell = -\tau_a^{\ell -}\ (\leq 0), \qquad a \in \hat{A}$$

Calculation of multipliers $\tau_a^{\ell-}$

Considering the dual problem of (`SUBP` ℓ), its solution will verify, for link $a=(i,j)\in A$:

$$c_a^\ell = \hat{u}_i^\ell - \hat{u}_j^\ell + \hat{\tau}_a^{\ell+} - \hat{\tau}_a^{\ell-}$$

In order to determine multipliers τ the following must be taken into account. Given a *master problem* solution for given decision variables $\bar{y}_a,\ a\in\hat{A}$, resolving `SUBP` ℓ would be carried out by eliminating links $a\in\hat{A}$ for which $\bar{y}_a=0$. Then, it is clear that:

- If $a=(i,j)\notin\hat{A}$, $\hat{\tau}_a^{\ell-}=0$.
- If $a=(i,j)\in\hat{A}$ and, additionally $\bar{y}_a>0$, then $\tau_a^{\ell-}=0$, whereas $\tau^{\ell+}$ may be null or positive. Always, $\hat{\tau}_a^{\ell+}=c_a^\ell-(\hat{u}_i^\ell-\hat{u}_j^\ell)$ will hold. In any case, it will be shown that $\tau_a^{\ell+}$ do not intervene at all.
- Si $a=(i,j)\in\hat{A}$ and, additionally $\bar{y}_a=0$:
  - If $c_a^\ell-(\hat{u}_i^\ell-\hat{u}_j^\ell)<0$ then: $\hat{\tau}_a^{\ell+}=0,\quad \hat{\tau}_a^{\ell-}=-c_a^\ell+(\hat{u}_i^\ell-\hat{u}_j^\ell)\,(>0)$. (in this case if link $a\in\hat{A}$ is added its flow would be >0 for commodity ℓ)
  - Si $c_a^\ell-(\hat{u}_i^\ell-\hat{u}_j^\ell)\geq 0$ then: $\hat{\tau}_a^{\ell-}=0,\quad \hat{\tau}_a^{\ell+}=c_a^\ell-(\hat{u}_i^\ell-\hat{u}_j^\ell)\,(>0)$. (in this case, if link $a\in\hat{A}$ is added its flow would be 0 for commodity ℓ)

  Then,in case of $\bar{y}_a=0$ the expression for $\hat{\tau}_a^{\ell-}$ would be:

$$\hat{\tau}_a^{\ell-} = max\{\ 0,\ (\hat{u}_i^\ell-\hat{u}_j^\ell)-c_a^\ell\ \}$$

Master Problem's structure:

taking into account the compact formulation in (1'), (2') and in case that the subproblem will never issue extreme directions, *master problem*'s structure at iteration M will be:

$$\begin{array}{rl} Min_{\ y,z} & z \\ s.t: & z\geq f^\top y+\sum_{\ell\in K}(\mathcal{T}^\ell+Fy)^\top\hat{\theta}^{\ell,s},\ s=1,2,3,...M \\ & y\in Y=\{0,1\}^{|\hat{A}|} \end{array}$$

Thus, the only thing that remains is to state the expression of the term $(\mathcal{T}^\ell+Fy)^\top\hat{\theta}^\ell$:

(super-index s omitted)

$$(\mathcal{T}^\ell+Fy)^\top\hat{\theta}^\ell=\left(\begin{pmatrix}t^\ell\\0\end{pmatrix}+\rho\begin{pmatrix}0\\y\end{pmatrix}\right)^\top\begin{pmatrix}\hat{u}^\ell\\ \hat{\tau}^\ell\end{pmatrix}=t^{\ell\top}\hat{u}^\ell-\rho\sum_{\substack{a\in\hat{A}\\ \bar{y}_a=0}}\hat{\tau}_a^{\ell-}y_a$$

Finally, it must be remarked that when *subproblem* is solved in its primal form (`PRIMAL`), duality can be used so in order to evaluate $c^{\ell\top}\hat{x}^\ell=t^{\ell\top}\hat{u}^\ell$