Benders Decomposition

See Nemhauser and Wolsey, section II.3.7 and III.5.4, for more information.

Benders Decomposition reduces a mixed integer optimization problem with \( p \) continuous variables and \( n \) integer variables to one with just one continuous variable, and still \( n \) integer variables, but typically with an enormous number of constraints.

The initial problem is

\[
\begin{align*}
\max_{x,y} & \quad c^T x + h^T y \\
\text{subject to} & \quad Ax + Gy \leq b \\
& \quad x \in X \subseteq \mathbb{Z}_+^n, \ y \in \mathbb{R}_+^p
\end{align*}
\]

where \( b \in \mathbb{R}_+^m \) and \( c, h, A, \) and \( G \) are all dimensioned appropriately.

For each possible choice of \( \bar{x} \in X \), we could find the best choice for \( y \) by solving a linear program, so we could regard \( y \) as a function of \( x \). We can then replace the contribution of \( y \) to the objective by a scalar variable representing the value of the best choice for a given \( \bar{x} \). We start out with a crude approximation to this value, and then generate a sequence of dual solutions to tighten up the approximation.

Let \( x \in X \). We denote the value of the best choice for \( y \) by \( z_{LP}(x) \). The original mixed integer program can then be written as a nonconvex problem in the integer variables:

\[
\max_{x \in X} c^T x + z_{LP}(x)
\]

The function \( z_{LP}(x) \) is a concave piecewise linear function. We have

\[
z_{LP}(x) := \max_y h^T y \quad \text{subject to} \quad Gy \leq b - Ax, \ y \in \mathbb{R}_+^p
\]

By LP duality, we can also write

\[
z_{LP}(x) = \min_u (b - Ax)^T u \quad \text{subject to} \quad G^T u \geq h, \ u \in \mathbb{R}_+^m
\]

Note that the feasible region \( Q = \{u \in \mathbb{R}_+^m : G^T u \geq h\} \) for the dual problem does not depend on \( x \). We denote the extreme points and extreme rays of \( Q \) as \( K \) and \( J \) respectively:

\[
\begin{align*}
\text{extreme points:} & \quad u^k, \ k \in K \\
\text{extreme rays:} & \quad r^j, \ j \in J
\end{align*}
\]

If the inner product between \( (b - Ax) \) and any ray \( r^j \) is negative then \( z_{LP}(x) = -\infty \). Equivalently, in this situation, problem (2) is infeasible, so \( x \) does not allow a feasible solution to the original problem (1). Thus, we have the valid constraints

\[
(b - Ax)^T r^j \geq 0 \quad \text{for} \ j \in J
\]
that must be satisfied by any \( x \) that is feasible in (1).

If \( x \) satisfies (4) then the value of \( z_{LP}(x) \) is given by

\[
z_{LP}(x) = \min_{k \in K} (b - Ax)^T u^k.
\]

Thus, problem (1) can be written equivalently as

\[
\max_{x,t} \quad c^T x + t
\]
subject to \((b - Ax)^T u^k \geq t\) for \( k \in K \)
\((b - Ax)^T r^j \geq 0\) for \( j \in J \)
\(t \in \mathbb{R}, x \in X\) \hspace{1cm} (5)

This problem has fewer variables than the original formulation (1), but it may have a huge number of constraints. Thus, these constraints are generated as needed, as cutting planes.

Let \( \hat{K} \subseteq K \) and \( \hat{J} \subseteq J \) denote the current known extreme points and extreme rays of \( Q \), respectively. The current relaxation of (1) and (5) is then

\[
\max_{x,t} \quad c^T x + t
\]
subject to \((b - Ax)^T u^k \geq t\) for \( k \in \hat{K} \)
\((b - Ax)^T r^j \geq 0\) for \( j \in \hat{J} \)
\(t \in \mathbb{R}, x \in X\) \hspace{1cm} (6)

The scalar variable \( t \) represents an estimate of \( z_{LP}(x) \).

The algorithm can then be written:

1. Determine (possibly empty) initial sets \( \hat{K} \) of extreme points and \( \hat{J} \) of extreme rays of \( Q \).
2. Solve problem (6), giving solution \( \bar{x} \) and corresponding \( \bar{t} \).
3. Determine \( z_{LP}(\bar{x}) \) by solving (3).
4. If \( z_{LP}(\bar{x}) = -\infty \), an extreme ray of \( Q \) has been found. Add the extreme ray to \( \hat{J} \) and return to Step 2.
5. If \( z_{LP}(\bar{x}) < \bar{t} \) and finite, an extreme point of \( Q \) has been found. Add the extreme point to \( \hat{K} \) and return to Step 2.
6. If \( z_{LP}(\bar{x}) = \bar{t} \) then \( \bar{x} \) solves (1), with optimal \( y \) equal to the solution to (2) with \( x = \bar{x} \).