Linear Programming

Chongnan Li
chongnanli1997@hotmail.com

Additional content 1: Dantzig-Wolfe Decomposition

Outline:

● Introduction

In this file, we consider methods for solving linear programs that exhibit **special structure**. In particular, we consider linear programs that are in **block angular form** and develop the Dantzig-Wolfe decomposition method for solving such problems.

● Decomposition for Block Angular Linear Programs

Consider a linear program of the form:

$$\begin{array}{ll} \text{Minimize} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array}$$

where $\boldsymbol{A}$ is an $m \times n$ matrix, and suppose that the constraint matrix is of the form:

$$\boldsymbol{A} = \begin{bmatrix} L_1 & L_2 & \cdots & L_K \\ A_1 & & & \\ & A_2 & & \vdots \\ & & \ddots & \\ & & & A_K \end{bmatrix}$$

where L_k is a submatrix with dimension $m_L \times n_k$ for $k = 1, \cdots, K$ and A_k is a submatrix of dimension $m_k \times n_k$ for $k = 1, \cdots, K$ such that $\sum_{k=1}^{K} n_k = n$ and $m_L + \sum_{k=1}^{K} m_k = m$. Such a form for $\boldsymbol{A}$ is called **block angular**.

Let $\boldsymbol{x}^k$, $\boldsymbol{c}^k$ and $\boldsymbol{b}^k$ be corresponding vectors such that

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}^1 \\ \vdots \\ \boldsymbol{x}^K \end{bmatrix}, \ \boldsymbol{c} = \begin{bmatrix} \boldsymbol{c}^1 \\ \vdots \\ \boldsymbol{c}^K \end{bmatrix} \text{ and } \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}^0 \\ \boldsymbol{b}^1 \\ \vdots \\ \boldsymbol{b}^K \end{bmatrix}, \ \text{so that } \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}.$$

Then, the linear program can be written as

$$\begin{array}{lllllll}
\text{Minimize} & (\boldsymbol{c}^1)^T\boldsymbol{x}^1 + (\boldsymbol{c}^2)^T\boldsymbol{x}^2 + \cdots + (\boldsymbol{c}^K)^T\boldsymbol{x}^K & \\
\text{subject to} & L_1\boldsymbol{x}^1 + L_2\boldsymbol{x}^2 + \cdots + L_K\boldsymbol{x}^K \le \boldsymbol{b}^0 & \\
& A_1\boldsymbol{x}^1 \qquad\qquad\qquad\qquad\quad \le \boldsymbol{b}^1 & \\
& \qquad\quad A_2\boldsymbol{x}^2 \qquad\qquad\qquad \le \boldsymbol{b}^2 & \\
& \qquad\qquad\qquad\quad \ddots \qquad\qquad \vdots\ \vdots & \\
& \qquad\qquad\qquad\qquad\quad A_K\boldsymbol{x}^K \ \le \boldsymbol{b}^K & \\
& \boldsymbol{x}^1 \ge \boldsymbol{0},\ \boldsymbol{x}^2 \ge \boldsymbol{0},\ \cdots,\ \boldsymbol{x}^K \ge \boldsymbol{0} &
\end{array}$$

or more compactly as

$$\begin{array}{ll}
\text{Minimize} & \displaystyle\sum_{k=1}^{K} (\boldsymbol{c}^k)^T\boldsymbol{x}^k \\
\text{subject to} & \displaystyle\sum_{k=1}^{K} L_k\boldsymbol{x}^k \le \boldsymbol{b}^0 \\
& A_k\boldsymbol{x}^k \le \boldsymbol{b}^k \qquad k = 1,\cdots,K \\
& \boldsymbol{x}^k \ge \boldsymbol{0} \qquad\quad\ k = 1,\cdots,K
\end{array}$$

The constraints $\displaystyle\sum_{k=1}^{K} L_k\boldsymbol{x}^k \le \boldsymbol{b}^0$ are called **coupling** or **linking** constraints (the number of such constraints is m_L) since without them, the problem decomposes into K independent subproblems where the kth subproblem is:

$$(SP_{\mathrm{k}})\begin{cases} \text{Minimize} & (\boldsymbol{c}^k)^T\boldsymbol{x}^k \\ \text{subject to} & A_k\boldsymbol{x}^k \le \boldsymbol{b}^k \\ & \boldsymbol{x}^k \ge \boldsymbol{0} \end{cases}$$

where each **sub-problem** is a linear program.

The idea in decomposition is to exploit the structure of such a linear program so that the entire problem does not have to be solved **at once**, but instead solve problems that are smaller and usually more tractable. In Dantzig-Wolfe decomposition, the problem will be decomposed into a master problem that will be concerned with the linking constraints only, and into subproblems that result from the decoupling of the linking constraints.

The master problem (MP) is of the form **(no nonnegativity constraints!)**:

$$
(\text{MP}) \left\{ \begin{array}{ll} \text{Minimize} & \sum_{k=1}^{K} (\boldsymbol{c}^k)^T \boldsymbol{x}^k \\ \text{subject to} & \sum_{k=1}^{K} L_k \boldsymbol{x}^k \leq \boldsymbol{b}^0 \end{array} \right.
$$

- Master problem **reformulation**

It is not enough to have the decompositon into master and subproblems as above. To make the decomposition **effective**, the master problem needs reformulation.

The reformulation will enable the master problem and subproblems to **exchange information** regarding progress toward an optimal solution for the original linear program while **enabling each problem to be solved separately**.

The key to reformulation is the fact that the **subproblems** are all linear programs and so the feasible sets are **polyhedrons**. Recall from the **Resolution Theorem** that any feasible point of a linear program can be represented as a **convex combination of the extreme points and a non-negative linear combination of extreme directions of the feasible set**.

Let $P_k = \{\boldsymbol{x}^k \,|\, A_k \boldsymbol{x}^k \leq \boldsymbol{b}^k,\ \boldsymbol{x}^k \geq \boldsymbol{0}\}$ be the feasible set of subproblem (SP_k) and $\boldsymbol{v}_1^k, \boldsymbol{v}_2^k, \cdots, \boldsymbol{v}_{N_k}^k$ be the extreme points of P_k and $\boldsymbol{d}_1^k, \boldsymbol{d}_2^k, \cdots, \boldsymbol{d}_{l_k}^k$ be the extreme directions of P_k. Then, by the Resolution Theorem, any point $\boldsymbol{x}^k \in P_k$ can be expressed as

$$
\boldsymbol{x}^k = \sum_{i=1}^{N_k} \lambda_i^k \boldsymbol{v}_i^k + \sum_{j=1}^{l_k} \mu_j^k \boldsymbol{d}_j^k
$$

where $\sum_{i=1}^{N_k} \lambda_i^k = 1$, $\lambda_i^k \geq 0$ for $i = 1, \cdots, N_k$ and $\mu_j^k \geq 0$ for $j = 1, \cdots, l_k$.

Substituting this representation into the master problem gives:

$$
\text{(MP)} \begin{cases}
\text{Minimize} & \displaystyle\sum_{k=1}^{K} (\boldsymbol{c}^k)^T \Big(\sum_{i=1}^{N_k} \lambda_i^k \boldsymbol{v}_i^k + \sum_{j=1}^{l_k} \mu_j^k \boldsymbol{d}_j^k \Big) \\
\text{subject to} & \displaystyle\sum_{k=1}^{K} L_k \Big(\sum_{i=1}^{N_k} \lambda_i^k \boldsymbol{v}_i^k + \sum_{j=1}^{l_k} \mu_j^k \boldsymbol{d}_j^k \Big) \le \boldsymbol{b}^0 \\
& \displaystyle\sum_{i=1}^{N_k} \lambda_i^k = 1 \qquad\qquad k = 1, \cdots, K \\
& \lambda_i^k \ge 0 \quad i = 1, \cdots, N_k \qquad k = 1, \cdots, K \\
& \mu_j^k \ge 0 \quad j = 1, \cdots, l_k \qquad k = 1, \cdots, K
\end{cases}
$$

After **simplification**, the master problem becomes:

$$
\text{(MP)} \begin{cases}
\text{Minimize} & \displaystyle\sum_{k=1}^{K} \sum_{i=1}^{N_k} \lambda_i^k (\boldsymbol{c}^k)^T (\boldsymbol{v}_i^k) + \sum_{k=1}^{K} \sum_{j=1}^{l_k} \mu_j^k (\boldsymbol{c}^k)^T (\boldsymbol{d}_j^k) \\
\text{subject to} & \displaystyle\sum_{k=1}^{K} \sum_{i=1}^{N_k} \lambda_i^k (L_k \boldsymbol{v}_i^k) + \sum_{k=1}^{K} \sum_{j=1}^{l_k} \mu_j^k (L_k \boldsymbol{d}_j^k) \le \boldsymbol{b}^0 \\
& \displaystyle\sum_{i=1}^{N_k} \lambda_i^k = 1 \qquad\qquad k = 1, \cdots, K \\
& \lambda_i^k \ge 0 \quad i = 1, \cdots, N_k \qquad k = 1, \cdots, K \\
& \mu_j^k \ge 0 \quad j = 1, \cdots, l_k \qquad k = 1, \cdots, K
\end{cases}
$$

The master problem now has as variables λ_i^k and μ_j^k. Corresponding to each **extreme point** $\boldsymbol{v}_i^k \in P_k$.

Let

$$f_i^k = (\boldsymbol{c}^k)^T (\boldsymbol{v}_i^k) \text{ and } \boldsymbol{q}_i^k = L_k \boldsymbol{v}_i^k.$$

And corresponding to each **extreme direction** $\boldsymbol{d}_j^k$ of P_k, let

$$f_j^{-k} = (\boldsymbol{c}^k)^T (\boldsymbol{d}_j^k) \text{ and } \boldsymbol{q}_j^{-k} = L_k \boldsymbol{d}_j^k.$$

Then the master program can be **reformulated** as:

$$
(\text{MP})\begin{cases}
\text{Minimize} & \displaystyle\sum_{k=1}^{K}\sum_{i=1}^{N_k}\lambda_i^k f_i^k + \sum_{k=1}^{K}\sum_{j=1}^{l_k}\mu_j^k f_j^{-k} \\
\text{subject to} & \displaystyle\sum_{k=1}^{K}\sum_{i=1}^{N_k}\lambda_i^k \boldsymbol{q}_i^k + \sum_{k=1}^{K}\sum_{j=1}^{l_k}\mu_j^k \boldsymbol{q}_j^{-k} \le \boldsymbol{b}^0 \\
& \displaystyle\sum_{i=1}^{N_k}\lambda_i^k = 1 \qquad\qquad k = 1,\cdots,K \\
& \lambda_i^k \ge 0 \quad i = 1,\cdots,N_k \qquad k = 1,\cdots,K \\
& \mu_j^k \ge 0 \quad j = 1,\cdots,l_k \qquad k = 1,\cdots,K
\end{cases}
$$

The constraints of type $\sum_{i=1}^{N_k}\lambda_i^k = 1$ are called **convexity constraints** associated with **subproblem** (SP_k). The master problem can be more compactly represented as:

$$
(\text{MP})\begin{cases}
\text{Minimize} & \boldsymbol{f}_v^T\boldsymbol{\lambda} + \boldsymbol{f}_d^T\boldsymbol{\mu} \\
\text{subject to} & \boldsymbol{Q}_v\boldsymbol{\lambda} + \boldsymbol{Q}_d\boldsymbol{\mu} + \boldsymbol{s} = \boldsymbol{r} \\
& \boldsymbol{\lambda} \ge \boldsymbol{0},\ \boldsymbol{\mu} \ge \boldsymbol{0},\ \boldsymbol{s} \ge \boldsymbol{0}
\end{cases}
$$

where

$$\boldsymbol{\lambda} = (\lambda_1^1,\cdots,\lambda_{N_1}^1,\ \lambda_1^2,\cdots,\lambda_{N_2}^2,\ \cdots,\ \lambda_1^K,\cdots,\lambda_{N_K}^K)^T$$

$$\boldsymbol{\mu} = (\mu_1^1,\cdots,\mu_{l_1}^1,\ \mu_1^2,\cdots,\mu_{l_2}^2,\ \cdots,\ \mu_1^K,\cdots,\mu_{l_K}^K)^T$$

$$\boldsymbol{f}_v = (f_1^1,\cdots,f_{N_1}^1,\ f_1^2,\cdots,f_{N_2}^2,\cdots,\ f_1^K,\cdots,f_{N_K}^K)^T$$

$$\boldsymbol{f}_d = (f_1^{-1},\cdots,f_{l_1}^{-1},\ f_1^{-2},\cdots,f_{l_2}^{-2},\ \cdots,\ f_1^{-K},\cdots,f_{l_K}^{-K})^T$$

$$\boldsymbol{r}^T = [(\boldsymbol{b}^0)^T,\boldsymbol{e}^T] = [(\boldsymbol{b}^0)^T,\underbrace{(1,\cdots,1)^T}_{K}]$$

$$\boldsymbol{s}^T = [(\underbrace{\boldsymbol{s}^0}_{m_L})^T,\underbrace{(0,\cdots,0)}_{K}{}^T]$$

where $\boldsymbol{e}$ is the vector of dimension K with all components equal to 1. $\boldsymbol{Q}_v$ is a matrix such that **the column associated with λ_i^k** is

$$\begin{bmatrix}\boldsymbol{q}_i^k \\ \boldsymbol{e}_k\end{bmatrix} = \begin{bmatrix}L_k\boldsymbol{v}_i^k \\ \boldsymbol{e}_k\end{bmatrix}$$

where $\boldsymbol{e}_k$ is the kth unit vector.

$\boldsymbol{Q}_d$ is a matrix such that **the column associated with μ_j^k** is

$$\begin{bmatrix} \boldsymbol{q}_j^{-k} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} L_k \boldsymbol{d}_j^k \\ \mathbf{0} \end{bmatrix}$$

The number of variables λ_i^k and μ_j^k can be **extremely large** for even moderately sized problems since feasible sets of subproblems (linear programs) can have an **extremely large number** of extreme points and extreme directions. In other words, there will be **many more columns than rows** in the reformulated master problem.

● **Restricted** Master Problem and the Revised Simplex Method

The key idea to get around this difficulty of handling **an extremely large number of variables** is to use the *revised simplex method* to solve the master problem. The major advantage is that it is not necessary to formulate the entire reformulated master problem since the vast majority of the variables will be zero (i.e., non-basic) at an optimal (basic feasible) solution**(because the number of rows is much less than that of columns).**

This motivates the construction of a smaller version of the master problem called the **restricted master problem** where only a small subset of the variables λ_i^k and μ_j^k are included corresponding to a current basic feasible solution, and the remaining variables are non-basic, i.e., set to zero. If the reduced costs of the basic feasible solution are non-negative, then the revised simplex method stops with the optimal solution, else some non-basic variable with negative reduced cost is selected to enter the basis.

Given a curent basis $\boldsymbol{B}$ for the **restricted master problem**, let $\boldsymbol{\pi}^T = \boldsymbol{f}_B^T \boldsymbol{B}^{-1}$ where $\boldsymbol{f}_B$ is a vector consisting of the quantities f_i^k and f_j^{-k} associated with variables λ_i^k and μ_j^k, which are basic. We assume that the components of $\boldsymbol{\pi}$ are arranged so that

$$\boldsymbol{\pi} = \begin{bmatrix} \boldsymbol{\pi}^1 \\ \pi_1^2 \\ \vdots \\ \pi_K^2 \end{bmatrix}$$

where $\boldsymbol{\pi}^1$ are the dual variables associated with the **linking constraints** and π_i^2 is the dual variable associated with the **convexity constraint of subproblem** i in the restricted master problem.

Then, the reduced cost corresponding to a **non-basic variable** λ_i^k is of the form

$$r_i^k = f_i^k - \boldsymbol{\pi}^T \begin{bmatrix} \boldsymbol{q}_i^k \\ \boldsymbol{e}_k \end{bmatrix} = (\boldsymbol{c}^k)^T(\boldsymbol{v}_i^k) - (\boldsymbol{\pi}^1)^T L_k \boldsymbol{v}_i^k - \pi_k^2$$

and the reduced cost corresponding to a **non-basic variable** μ_j^k is of the form

$$r_j^{-k} = f_j^{-k} - \boldsymbol{\pi}^T \begin{bmatrix} \boldsymbol{q}_j^{-k} \\ \mathbf{0} \end{bmatrix} = (\boldsymbol{c}^k)^T(\boldsymbol{d}_j^k) - (\boldsymbol{\pi}^1)^T L_k \boldsymbol{d}_j^k$$

There will be a **considerable number of non-basic** variables, but fortunately one **does not have to compute all** of the reduced costs. In fact, it will suffice to determine **only the minimal** reduced cost among all of the non-basic variables. To this end, we let

$$r_{\min} = \min_{k \in \{1,\cdots,K\}} \left\{ \min_{i \in \{1,\cdots,N_k\}} \{r_i^k\} \right\}$$

or

$$r_{\min} = \min_{k \in \{1,\cdots,K\}} \left\{ \min_{i \in \{1,\cdots,N_k\}} \{(\boldsymbol{c}^k)^T(\boldsymbol{v}_i^k) - (\boldsymbol{\pi}^1)^T L_k \boldsymbol{v}_i^k - \pi_k^2\} \right\}$$

Let $r_*^k = \min_{i \in \{1,\cdots,N_k\}} \{r_i^k\}$. **Then r_*^k is equivalent to the optimal objective function of the subproblem (SP_k)** :

$$\begin{aligned} &\text{Minimize} && \sigma_k = \left((\boldsymbol{c}^k)^T - (\boldsymbol{\pi}^1)^T L_k\right) \boldsymbol{x}^k \\ &\text{subject to} && A_k \boldsymbol{x}^k \le \boldsymbol{b}^k \\ &&& \boldsymbol{x}^k \ge \boldsymbol{0} \end{aligned}$$

Since the term π_k^2 is fixed, it can be removed from the objective function. We assume that the **revised simplex method will be used to solve the subproblem**, so if the subproblem is bounded, an optimal extreme point $\boldsymbol{x}^k$ will be generated and **$\boldsymbol{x}^k$ will be one of the extreme points $\boldsymbol{v}_i^k$.**

Let the **optimal extreme point** of subproblem (SP_k) be $\boldsymbol{v}_{i*}^k$ for some index $i* \in \{1,\cdots,N_k\}$, and let σ_k^* denote the optimal objective function value of (SP_k). Then $r_*^k = \sigma_k^* - \pi_k^2$.

There are **three possibilities** in solving the subproblems (SP_k) in attempting to generate $r_{\min}$.

(1)If *all subproblems are bounded* and $r_{\min} = \min_{k \in \{1,\cdots,K\}} \{r_*^k\} < 0$, then let t be the index k such that $r_{\min} = r_*^t$. The column $\begin{bmatrix} \boldsymbol{q}_{i*}^t \\ \boldsymbol{e}_t \end{bmatrix} = \begin{bmatrix} L_t \boldsymbol{v}_{i*}^t \\ \boldsymbol{e}_t \end{bmatrix}$ associated with the optimal **extreme point** $\boldsymbol{v}_{i*}^t$ of subproblem (SP_t) that achieved $r_{\min} = r_*^t$ is entered into the basis $\boldsymbol{B}$.

(2)If *all subproblems are bounded* and $r_{\min} = \min_{k \in \{1,\cdots,K\}} \{r_*^k\} \ge 0$, then the current basis $\boldsymbol{B}$ is optimal.

(3)If there is *at least one subproblem that is unbounded*, then let s be the index k of such an **unbounded subproblem** (SP_s). The revised simplex method will return an **extreme direction** $\boldsymbol{d}_{j*}^s$ for some $j* \in \{1,\cdots,l_s\}$ associated with (SP_s) such that $\left((\boldsymbol{c}^s)^T - (\boldsymbol{\pi}^1)^T L_s\right) \boldsymbol{d}_{j*}^s < 0$, and so the column $\begin{bmatrix} \boldsymbol{q}_{j*}^{-s} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} L_s \boldsymbol{d}_{j*}^s \\ \boldsymbol{0} \end{bmatrix}$ associated with μ_{j*}^s (the multiplier of $\boldsymbol{d}_{j*}^s$) can enter the basis $\boldsymbol{B}$.

● **Price-Directed Decomposition**

The Dantzig-Wolfe decomposition can be interpreted as a **price-directed** decomposition in that the restricted master problem **generates prices in the form of dual values $\boldsymbol{\pi}$** and these **prices are sent to the subproblems (SP_k) to form the coefficients of the objective function σ_k**. The subproblems then **compute optimal solutions $\boldsymbol{x}^k$** or **generate extreme directions** and send back up to the **master problem** the appropriate column to possibly enter the basis. The process iterates until an optimal solution is found or the problem is declared unbounded.

● **Key Steps of Dantzig-Wolfe Decomposition**

We now provide the detailed steps of the Dantzig-Wolfe Decomposition.

Step 0: (Initialization)

Generate an initial basis $\boldsymbol{B}$ for the master problem.
Let $\boldsymbol{x}_B$ be the basic variables and $\bar{B}$ the index set of the basic variables and set all other variables to non-basic(zero) to get the restricted master problem. Go to Step 1.

Step 1: (Simplex Multiplier Generation)

Solve for $\boldsymbol{\pi}$ in the linear system $\boldsymbol{B}^T\boldsymbol{\pi} = \boldsymbol{f}_B$. Go to Step 2.

Step 2: (Optimality Check)

For each $k = 1,\cdots,K$, solve (SP_k), i.e.,

$$\begin{aligned} \text{Minimize} \quad & \sigma_k = \left((\boldsymbol{c}^k)^T - (\boldsymbol{\pi}^1)^T L_k\right)\boldsymbol{x}^k \\ \text{subject to} \quad & A_k\boldsymbol{x}^k \leq \boldsymbol{b}^k \\ & \boldsymbol{x}^k \geq \boldsymbol{0} \end{aligned}$$

using the revised simplex method. If (SP_k) is unbounded, then go to Step 3, else let $\boldsymbol{x}^k = \boldsymbol{v}^k_{i*}$ denote the optimal basic feasible solution and compute $r^k_* = \sigma^*_k - \pi^2_k$.
If $r_{\min} = \min\limits_{k\in\{1,\cdots,K\}} \{r^k_*\} \geq 0$, then STOP, the current basis $\boldsymbol{B}$ is optimal;
else go to Step 3.

Step 3: (Column Generation)

If all subproblems (SP_k) are bounded and $r_{\min} = \min\limits_{k\in\{1,\cdots,K\}} \{r^k_*\} < 0$, then let t be the index of k in (SP_k) such that $r_{\min} = r^t_*$. Let $\bar{\boldsymbol{a}} = \begin{bmatrix} \boldsymbol{q}^t_{i*} \\ \boldsymbol{e}_t \end{bmatrix} = \begin{bmatrix} L_t\boldsymbol{v}^t_{i*} \\ \boldsymbol{e}_t \end{bmatrix}$ where $\boldsymbol{v}^t_{i*}$ is the optimal extreme point of (SP_t) and go to Step 4. Else there is a subproblem (SP_s) that is unbounded, and so an extreme direction $\boldsymbol{d}^s_{j*}$ will be generated such that $\left((\boldsymbol{c}^s)^T - (\boldsymbol{\pi}^1)^T L_s\right)\boldsymbol{d}^s_{j*} < 0$, and so let $\bar{\boldsymbol{a}} = \begin{bmatrix} \boldsymbol{q}^{-s}_{j*} \\ \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} L_s\boldsymbol{d}^s_{j*} \\ \boldsymbol{0} \end{bmatrix}$ and go to Step 4.

Step 4: (Descent Direction Generation)
Solve for $\boldsymbol{d}$ in the linear system $\boldsymbol{B}\boldsymbol{d} = -\bar{\boldsymbol{a}}$. If $\boldsymbol{d} \geq \boldsymbol{0}$, then the linear program is unbounded, STOP. Else go to Step 5.

Step 5: (Step Length Generation)
Compute the step length $\alpha = \min_{l \in \bar{B}}\{-\frac{x_l}{d_l} \mid d_l < 0\}$ (the minimum ratio test). Let l^* be the index of the basic variable that attains the minimum ratio α. Go to Step 6.

Step 6: (Update Basic Variables)
Now let $\boldsymbol{x}_B := \boldsymbol{x}_B + \alpha\boldsymbol{d}$. Go to Step 7.

Step 7: (Basis Update)
Let $\boldsymbol{B}_{l^*}$ be the column in $\boldsymbol{B}$ associated with the leaving basic variable x_{l^*}. Update the basis matrix $\boldsymbol{B}$ by removing $\boldsymbol{B}_{l^*}$ and adding the column $\bar{\boldsymbol{a}}$ and update $\bar{B}$. Go to Step 1.

● Economic Interpretation

The Dantzig-Wolfe decomposition algorithm has the following economic interpretation, which supports the view that it is a **price-directed decomposition** as described earlier.

The original linear program with block angular structure represents a **company with K subdivisions (subproblems)** where **subdivision k independently produces its own set of products according to its set of constraints $A_k\boldsymbol{x}^k \leq \boldsymbol{b}^k$**. But all subdivisions require the use of a limited set of **common resources,** which give rise to the **linking constraints**. The company wishes to minimize the cost of production over **all products** from the K subdivisions.

The master problem represents a company-wide supervisor that manages the use of the **common resources**. At an iteration of the Dantzig-Wolfe decomposition, the restricted master problem generates a **master production plan** represented by the basis $\boldsymbol{B}$ based on the **production plan (proposals) sent by the subdivisions.** The basis represents that **fraction or weight** of each proposal in the master plan.

Then, the supervisor is responsible for the computation of the vector $\boldsymbol{\pi}^1$ which represents the calculation of the prices for the common resources where $-\pi_i^1$ is the price for consuming **a unit of common resource i**. These **(marginal)** prices reflect demand for the common resources by the master plan, are announced to all of the subdivisions, and are used by each subdivision in the coefficients of its objective function in constructing its optimal production plan (proposal). Recall that the objective of subdivision k is to minimize

$$\sigma_k = \left((\boldsymbol{c}^k)^T - (\boldsymbol{\pi}^1)^T L_k\right)\boldsymbol{x}^k$$

where $(\boldsymbol{c}^k)^T\boldsymbol{x}^k$ is the original objective function of the subdivision and represents its **own specific (i.e., not related to other subdivisions) costs**, and $L_k\boldsymbol{x}^k$, is the **quantity of common resources** consumed by the production plan (proposal) $\boldsymbol{x}^k$. Thus, the total **costs** of using common resources by subdivision k is represented by

$-(\boldsymbol{\pi}^1)^T L_k \boldsymbol{x}^k$, and so the more it uses the common resources the worse the overall objective will be.

The proposal that gives the **greatest** promise in reducing costs for the company (i.e., that proposal that achieves $r_{\min} < 0$) is selected by the supervisor. If all proposals are such that $r_{\min} \geq 0$, then there are **no further cost savings** possible and the current production plan is optimal. **However**, if a proposal is selected and a new production plan is generated by the supervisor since the weights of the current proposals must be adjusted (this is accomplished by the updating the basic variables) so that the capacities of the common resources **are not violated** with the introduction of a new proposal. A **new set of prices** are generated and the process iterates.

Thus, the Dantzig-Wolfe decompositon represents a decentralized mechanism for resource allocation as the decision making for actual production lies within the subdivisions and the coordination is accomplished through the **prices** set by the supervisor which reflect supply and demand. The prices serve to guide the subdivisions to a production plan that is systemwide (company-wide) cost optimal, which represents an equilibrium where supply and demand are balanced.

- Initialization

To start the Dantzig-Wolfe decomposition, the master problem requires an initial basic feasible solution. The strategy is to develop an **auxiliary problem** similar to a Phase I approach for the revised simplex method.

First, an extreme point $\boldsymbol{x}^k = \boldsymbol{v}_1^k$ is generated for each subproblem (SP_k) using the Phase I approach for the revised simplex method. If any of the subproblems do not have a feasible extreme point, then the original problem is infeasible. Even if all subproblems admit an extreme point, **it might be the case that the linking constraints might be violated.**

Like the Phase I procedure for the simplex method, artificial variables $\boldsymbol{x}_a$ can be added to the linking constraints and these artificial variables are then minimized. So the **auxiliary problem** for the master problem is

$$
\begin{aligned}
\text{Minimize} \quad & \boldsymbol{e}^T \boldsymbol{x}_a \\
\text{subject to} \quad & \sum_{k=1}^{K} \sum_{i=1}^{N_k} \lambda_i^k (L_k \boldsymbol{v}_i^k) + \sum_{k=1}^{K} \sum_{j=1}^{l_k} \mu_j^k (L_k \boldsymbol{d}_j^k) + \boldsymbol{x}_a = \boldsymbol{b}^0 \\
& \sum_{i=1}^{N_k} \lambda_i^k = 1, \quad k = 1,\cdots,K \\
& \lambda_i^k \geq 0, \quad i = 1,\cdots,N_k \quad k = 1,\cdots,K \\
& \mu_j^k \geq 0, \quad j = 1,\cdots,l_k \quad k = 1,\cdots,K \\
& \boldsymbol{x}_a \geq \boldsymbol{0}
\end{aligned}
$$

An initial basic feasible solution for the auxiliary problem is to let $\lambda_1^k = 1$ for all $k = 1,\cdots,K$; $\lambda_i^k = 0$ for $i \neq 1$, $\mu_j^k = 0$ for all k and j, and

$$\boldsymbol{x}_a = \boldsymbol{b}^0 - \sum_{k=1}^{K} \lambda_1^k (L_k \boldsymbol{v}_1^k) = \boldsymbol{b}^0 - \sum_{k=1}^{K} L_k \boldsymbol{v}_1^k$$

Recall that the points $\boldsymbol{v}_1^k$ are generated earlier by the Phase I method applied to (SP_k). Therefore, if $\boldsymbol{e}^T \boldsymbol{x}_a > 0$, then the master problem is infeasible, else $\boldsymbol{e}^T \boldsymbol{x}_a = 0$ and the optimal solution to the auxiliary problem will provide an initial basic feasible solution for the master problem.

NOTE: if $\boldsymbol{b} \geq \boldsymbol{0}$, then the initialization for the problem would be **much easier.** Because we can set $\boldsymbol{v}_1^k$ to the original point for all $k = 1, \cdots, K$, set $\lambda_1^k = 1$, for all $k = 1, \cdots, K$ and let $\boldsymbol{s} = \boldsymbol{b}^0$ as the initial basic solution.

● Bounds on Optimal Cost

It is found through computational experiments that the Dantzig-Wolfe decomposition **can often take too much time** before termination in solving very large problem instances. However, the method can be stopped before optimality has achieved and one can evaluate to some extent how close the current basic feasible solution is from optimal. An objective function value based on a feasible solution obtained before optimality represents **an upper bound** (assuming a minimization problem) on the optimal objective function value.

The idea is to now generate a **feasible solution of the dual of the restricted master problem**; the associated objective function value will give **a lower bound**. Thus, both upper and lower bounds can be obtained on the optimal cost.

Theorem 5.4

Suppose the master problem is *consistent* **(When $P \neq \phi$, LP is *consistent*.)** and *bounded* with optimal objective function value z^*. Let $\boldsymbol{x}_B$ be the basic variables from a feasible solution obtained from the *Dantzig-Wolfe decomposition* before termination and denote z as its correponding objective function value. Further assume that all subproblems (SP_k) $(\forall k = 1, \cdots, K)$ are *bounded* with optimal objective function value σ_k^*, then

$$\left(z + \sum_{k=1}^{K} r_*^k \right) = z + \sum_{k=1}^{K} (\sigma_k^* - \pi_k^2) \leq z^* \leq z$$

where π_k^2 is the dual variable associated with the convexity constraint of subproblem k in the master problem.

Before the proof of Theorem 5.4 is presented, we give the following lemma.

Lemma 5.5

Let $\boldsymbol{x}$ be a *non-optimal feasible* solution for the master problem that is generated during an intermediate step of the *Dantzig-Wolfe decomposition* and suppose that all subproblems are *bounded*. Then, a *feasible* solution for the *dual of the master problem* exists whose objective function value is equal to the objective function value of master problem at $\boldsymbol{x}$.

Proof:

Let z denote the objective function of the master problem at $\boldsymbol{x}$. Without loss of generality, we consider the case that there are only two subproblems and all complicating constraints are equality constraints, and so the master problem is

$$
\begin{array}{lll}
\text{Minimize} & \displaystyle\sum_{i=1}^{N_1} \lambda_i^1 f_i^1 + \sum_{i=1}^{N_2} \lambda_i^2 f_i^2 + \sum_{j=1}^{l_1} \mu_j^1 f_j^{-1} + \sum_{j=1}^{l_2} \mu_j^2 f_j^{-2} & \\
\text{subject to} & \displaystyle\sum_{i=1}^{N_1} \lambda_i^1 L_1 \boldsymbol{v}_i^1 + \sum_{i=1}^{N_2} \lambda_i^2 L_2 \boldsymbol{v}_i^2 + \sum_{j=1}^{l_1} \mu_j^1 L_1 \boldsymbol{d}_j^1 + \sum_{j=1}^{l_2} \mu_j^2 L_2 \boldsymbol{d}_j^2 = \boldsymbol{b}^0 & \\
& \displaystyle\sum_{i=1}^{N_1} \lambda_i^1 = 1 & \\
& \qquad\qquad \displaystyle\sum_{i=1}^{N_2} \lambda_i^2 = 1 & \\
& \lambda_i^1 \geq 0 & i = 1, \cdots, N_1 \\
& \qquad\qquad \lambda_i^2 \geq 0 & i = 1, \cdots, N_2 \\
& \mu_j^1 \geq 0 & j = 1, \cdots, l_1 \\
& \mu_j^2 \geq 0 & j = 1, \cdots, l_2
\end{array}
$$

The dual of the master problem is then

$$
\begin{array}{lll}
\text{Maximize} & (\boldsymbol{\pi}^1)^T \boldsymbol{b}^0 + \pi_1^2 + \pi_2^2 & \\
\text{subject to} & (\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_i^1 + \pi_1^2 \leq f_i^1, & \forall \boldsymbol{v}_i^1 \in P_1 \\
& (\boldsymbol{\pi}^1)^T L_1 \boldsymbol{d}_j^1 \leq f_j^{-1}, & \forall \boldsymbol{d}_j^1 \in P_1 \\
& (\boldsymbol{\pi}^1)^T L_2 \boldsymbol{v}_i^2 + \pi_2^2 \leq f_i^2, & \forall \boldsymbol{v}_i^2 \in P_2 \\
& (\boldsymbol{\pi}^1)^T L_2 \boldsymbol{d}_j^2 \leq f_j^{-2}, & \forall \boldsymbol{d}_j^2 \in P_2
\end{array}
$$

Assuming that the master problem is solved using the revised simplex method, there is a dual solution (simplex multipliers) $\boldsymbol{\pi} = [(\boldsymbol{\pi}^1)^T \ \pi_1^2 \ \pi_2^2]^T$ such that the objective function value of the dual $(\boldsymbol{\pi}^1)^T\boldsymbol{b}^0 + \pi_1^2 + \pi_2^2 = z$. However, $\boldsymbol{\pi}$ is infeasible except at optimality of the revised simplex method. To obtain a feasible dual solution, the boundness of the subproblems are exploited.

The first subproblem (SP_1) is

$$\begin{aligned} \text{Minimize} \quad & \sigma_1 = \left((\boldsymbol{c}^1)^T - (\boldsymbol{\pi}^1)^T L_1\right)\boldsymbol{x}^1 \\ \text{subject to} \quad & A_1\boldsymbol{x}^1 \le \boldsymbol{b}^1 \\ & \boldsymbol{x}^1 \ge \boldsymbol{0} \end{aligned}$$

The optimal solution is some extreme point $\boldsymbol{v}_{i*}^1 \in P_1$ with corresponding finite objective function value

$$\sigma_1^* = \left((\boldsymbol{c}^1)^T - (\boldsymbol{\pi}^1)^T L_1\right)\boldsymbol{v}_{i*}^1 = (\boldsymbol{c}^1)^T\boldsymbol{v}_{i*}^1 - (\boldsymbol{\pi}^1)^T L_1\boldsymbol{v}_{i*}^1.$$

Furthermore, there is no directions in P_1 for which (SP_1) is bounded, and so we have for all directions $\boldsymbol{d}_j^1 \in P_1$

$$f_j^{-1} - (\boldsymbol{\pi}^1)^T L_1\boldsymbol{d}_j^1 = (\boldsymbol{c}^1)^T\boldsymbol{d}_j^1 - (\boldsymbol{\pi}^1)^T L_1\boldsymbol{d}_j^1 \ge 0.$$

This suggests that we can use the quantity σ_1^* instead of π_1^2 **since $\boldsymbol{\pi}^1$ and σ_1^* are feasible for the first constraints of the dual problem.** In a similar fashion, since the second subproblem (SP_2) is bounded, we can use

$$\sigma_2^* = \left((\boldsymbol{c}^2)^T - (\boldsymbol{\pi}^1)^T L_2\right)\boldsymbol{v}_{i*}^2 = (\boldsymbol{c}^2)^T\boldsymbol{v}_{i*}^2 - (\boldsymbol{\pi}^1)^T L_2\boldsymbol{v}_{i*}^2$$

in place of π_2^2. Therefore, the solution $\boldsymbol{\pi}^f = [\boldsymbol{\pi}^1\ \sigma_1^*\ \sigma_2^*]^T$ is feasible for the dual problem.

Q.E.D.

Additional explanation for "since $\boldsymbol{\pi}^1$ and σ_1^* are feasible for the first constraints of the dual problem":

The first constraints of the dual problem are:

$$\begin{aligned} (\boldsymbol{\pi}^1)^T L_1\boldsymbol{v}_i^1 + \pi_1^2 &\le f_i^1, \quad \forall \boldsymbol{v}_i^1 \in P_1 \\ (\boldsymbol{\pi}^1)^T L_1\boldsymbol{d}_j^1 &\le f_j^{-1}, \quad \forall \boldsymbol{d}_j^1 \in P_1 \end{aligned}$$

where

$$(\boldsymbol{\pi}^1)^T L_1\boldsymbol{d}_j^1 \le f_j^{-1}, \quad \forall \boldsymbol{d}_j^1 \in P_1$$

are satisfied because of

$$f_j^{-1} - (\boldsymbol{\pi}^1)^T L_1\boldsymbol{d}_j^1 = (\boldsymbol{c}^1)^T\boldsymbol{d}_j^1 - (\boldsymbol{\pi}^1)^T L_1\boldsymbol{d}_j^1 \ge 0, \forall \boldsymbol{d}_j^1 \in P_1.$$

Consider

$$(\boldsymbol{\pi}^1)^T L_1\boldsymbol{v}_i^1 + \pi_1^2 \le f_i^1, \quad \forall \boldsymbol{v}_i^1 \in P_1$$

Choose any $i' \in \{1, \cdots, N_1\}$ such that $\boldsymbol{v}_{i'}^1 \in P_1$.

$$(\boldsymbol{\pi}^1)^T L_1\boldsymbol{v}_{i'}^1 + \pi_1^2 \le f_{i'}^1$$

Recall that we use the quantity σ_1^* instead of π_1^2, thus we substitue $\sigma_1^* = \pi_1^2$ into the constraint:

$$\begin{aligned}
&(\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_{i'}^1 + \pi_1^2 \ \leq f_{i'}^1 \\
&\Updownarrow \\
&(\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_{i'}^1 + \sigma_1^* \ \leq f_{i'}^1 \\
&\Updownarrow \\
&(\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_{i'}^1 + (\boldsymbol{c}^1)^T \boldsymbol{v}_{i*}^1 - (\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_{i*}^1 \ \leq f_{i'}^1 \\
&\Updownarrow \\
&(\boldsymbol{c}^1)^T \boldsymbol{v}_{i*}^1 - (\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_{i*}^1 \ \leq f_{i'}^1 - (\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_{i'}^1 \\
&\Updownarrow \\
&(\boldsymbol{c}^1)^T \boldsymbol{v}_{i*}^1 - (\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_{i*}^1 \ \leq (\boldsymbol{c}^1)^T \boldsymbol{v}_{i'}^1 - (\boldsymbol{\pi}^1)^T L_1 \boldsymbol{v}_{i'}^1
\end{aligned}$$

Obviously, the last red constraint is correct since the optimal solution of the first subproblem is $\boldsymbol{v}_{i*}^1 \in P_1$.

Proof of Theorem 5.4:

Now the objective function value of the master problem at $\boldsymbol{x}$ is

$$z = (\boldsymbol{\pi}^1)^T \boldsymbol{b}^0 + \pi_1^2 + \pi_2^2$$

where $\boldsymbol{\pi} = [(\boldsymbol{\pi}^1)^T \ \pi_1^2 \ \pi_2^2]^T$ is the vector of simplex multipliers associated with $\boldsymbol{x}$. By Lemma 5.5, the solution $\boldsymbol{\pi}^f = [(\boldsymbol{\pi}^1)^T \ \sigma_1^* \ \sigma_2^*]^T$ is dual feasible.

Thus, by weak duality of linear programming we have

$$\begin{aligned}
z^* &\geq (\boldsymbol{\pi}^1)^T \boldsymbol{b}^0 + \sigma_1^* + \sigma_2^* \\
&= \left[(\boldsymbol{\pi}^1)^T \boldsymbol{b}^0 + \pi_1^2 + \pi_2^2\right] + \sigma_1^* - \pi_1^2 + \sigma_2^* - \pi_2^2 \\
&= z + \sum_{i=1}^{2} (\sigma_i^* - \pi_i^2)
\end{aligned}$$

Q.E.D.

(这一段我实在不想用英语说了，我来稍加解释一下这个定理的证明。弱对偶定理是说，一个原始解，一个对偶解，如果他俩都是可行，那么原始解算出的目标函数值一定 ≥ 对偶解算出的目标函数值。证明过程中，$\boldsymbol{\pi} = [(\boldsymbol{\pi}^1)^T \ \pi_1^2 \ \pi_2^2]^T$是根据 $\boldsymbol{B}^T\boldsymbol{\pi} = \boldsymbol{f}_B$ 算出的单纯形算子，我们知道此时 $\boldsymbol{B}$ 还不是最优解的基，因此尽管主问题和主问题的对偶问题满足目标函数值相等，但是单纯形算子，即对偶变量是不可行的。然后那个引理是想表达，嘿哥们，我们把π_1^2, π_2^2给换一下吧，换成此时子问题的最优目标函数值σ_1^*, σ_2^*，那么这时候，对偶变量就可行了。好了，你现在有个原始可行解、有个对偶可行解，得嘞，用弱对偶吧。剩下的不需再多说了，应该就明白了。)

Example 5.6

Consider the basic variables at the end of the second iteration in Example 5.2:

$$\boldsymbol{x}_B = \begin{bmatrix} s_1 \\ \lambda_2^1 \\ \lambda_1^1 \\ \lambda_2^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.5 \\ 0.5 \\ 1 \end{bmatrix}, \text{ then } z = \boldsymbol{f}_B^T \boldsymbol{x}_B = [0 \quad -6 \quad 0 \quad -9] \begin{bmatrix} 1 \\ 0.5 \\ 0.5 \\ 1 \end{bmatrix} = -12.$$

Also, $\sum_{k=1}^{2} (\sigma_k^* - \pi_k^2) = (\sigma_1^* - \pi_1^2) + (\sigma_2^* - \pi_2^2) = (-4 - 0) + (-\frac{10}{3} - (-3)) = -4\frac{1}{3}$,

so we have $-16\frac{1}{3} \le z^* \le -12$. Recall that for this problem $z^* = -14$.