Block Structure of Primal

\[
\begin{bmatrix}
A \\
T_1 & W \\
T_2 & W \\
\vdots \\
T_K & W
\end{bmatrix}
\]

L-Shaped method: ignore constraints of future stages
Block Structure of Dual

Dantzig-Wolfe decomposition: ignore variables
1 Algorithm Description [Infanger, Bertsimas]

2 Examples [Bertsimas]

3 Application of Dantzig-Wolfe in Stochastic Programming [BL, §5.5]
   - Reformulation of 2-Stage Stochastic Program
   - Algorithm Description

4 Application of Dantzig-Wolfe in Integer Programming [Vanderbeck]
   - Dantzig-Wolfe Reformulation
   - Relationship to Lagrange Relaxation
1. Algorithm Description [Infanger, Bertsimas]

2. Examples [Bertsimas]

3. Application of Dantzig-Wolfe in Stochastic Programming [BL, §5.5]
   - Reformulation of 2-Stage Stochastic Program
   - Algorithm Description

   - Dantzig-Wolfe Reformulation
   - Relationship to Lagrange Relaxation
The Problem

\[ z^* = \min c_1^T x_1 + c_2^T x_2 \]
\[ \text{s.t. } A_1 x_1 + A_2 x_2 = b \]
\[ B_1 x_1 = d_1 \]
\[ B_2 x_2 = d_2 \]
\[ x_1, x_2 \geq 0 \]

- \( x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2} \)
- \( b \in \mathbb{R}^{m}, d_1 \in \mathbb{R}^{m_1}, d_2 \in \mathbb{R}^{m_2} \)
- \( A_1 x_1 + A_2 x_2 = b \) are complicating/coupling constraints

Note: This will be the form of the dual of the 2-stage stochastic program (see slide 3)
Minkowski’s Representation Theorem

Every polyhedron $P$ can be represented in the form

$$P = \{ x \in \mathbb{R}^n : x = \sum_{j \in J} \lambda^j x^j + \sum_{r \in R} \mu^r w^r, \quad \sum_{j \in J} \lambda^j = 1, \lambda \in \mathbb{R}^{|J|}_+, \mu \in \mathbb{R}^{|R|}_+ \}$$

where

- $\{x^j, j \in J\}$ are the extreme points of $P$
- $\{w^r, r \in R\}$ are the extreme rays of $P$
Graphical Illustration of Minkowski’s Representation Theorem

\[ x^1, x^2, x^3: \text{extreme points} \]
\[ w^1, w^2: \text{extreme rays} \]
\[ x = \lambda x^2 + (1 - \lambda)x^3 + \mu w^2, \quad 0 \leq \lambda \leq 1, \quad \mu \geq 0 \]
The Feasible Region of the Subproblems

We represent $B_1 x_1 = d_1$ as

$$\sum_{j \in J_1} \lambda^j_1 x^j_1 + \sum_{r \in R_1} \mu^r_1 w^r_1, \lambda^j_1 \geq 0, \mu^r_1 \geq 0, \sum_{j \in J_1} \lambda^j_1 = 1$$

and $B_2 x_2 = d_2$ as

$$\sum_{j \in J_2} \lambda^j_2 x^j_2 + \sum_{r \in R_2} \mu^r_2 w^r_2, \lambda^j_2 \geq 0, \mu^r_2 \geq 0, \sum_{j \in J_2} \lambda^j_2 = 1$$
Transform the full master problem using

- \( x_1 = \sum_{j \in J_1} \lambda_1^j x_1^j + \sum_{r \in R_1} \mu_1^r w_1^r \)
- \( x_2 = \sum_{j \in J_2} \lambda_1^j x_2^j + \sum_{r \in R_2} \mu_2^r w_2^r \)

For example,

\[ A_1 x_1 + A_2 x_2 = b \]

becomes

\[ \sum_{j \in J_1} \lambda_1^j A_1 x_1^j + \sum_{r \in R_1} \mu_1^r A_1 w_1^r + \sum_{j \in J_2} \lambda_2^j A_2 x_2^j + \sum_{r \in R_2} \mu_2^r A_2 w_2^r = b \]
Applying Minkowski’s representation theorem we obtain:

\[ z = \min \sum_{j \in J_1} \lambda_j^1 c_1^T x_j^1 + \sum_{r \in R_1} \mu_r^1 c_1^T w_r^1 + \sum_{j \in J_2} \lambda_j^2 c_2^T x_j^2 + \sum_{r \in R_2} \mu_r^2 c_2^T w_r^2 \]

\[ \sum_{j \in J_1} \lambda_j^1 A_1 x_j^1 + \sum_{r \in R_1} \mu_r^1 A_1 w_r^1 + \sum_{j \in J_2} \lambda_j^2 A_2 x_j^2 + \sum_{r \in R_2} \mu_r^2 A_2 w_r^2 = b, \ (\pi) \]

\[ \sum_{j \in J_1} \lambda_j^1 = 1, \ (t_1) \]

\[ \sum_{j \in J_2} \lambda_j^2 = 1, \ (t_2) \]

\[ \lambda_j^1, \lambda_j^2, \mu_r^1, \mu_r^2 \geq 0 \]
This problem is equivalent to the original problem.

The decision variables are the weights of the extreme points \( (\lambda^i_1, \lambda^i_2) \) and weights of the extreme rays \( (\mu^r_1, \mu^r_2) \).

The number of decision variables can be enormous (trick: we will ignore most of them).

The number of constraints is smaller (we got rid of \( B_1x_1 = d_1, B_2x_2 = d_2 \)).
Columns in the New Formulation

Constraint matrix in the new formulation:

\[
\sum_{j \in J_1} \lambda_j^1 \begin{bmatrix} A_1 x_1^j \\ 1 \\ 0 \end{bmatrix} + \sum_{j \in J_2} \lambda_j^2 \begin{bmatrix} A_2 x_2^j \\ 0 \\ 1 \end{bmatrix} + \sum_{r \in R_1} \mu_r^1 \begin{bmatrix} A_1 w_1^r \\ 0 \\ 0 \end{bmatrix} + \sum_{r \in R_2} \mu_r^2 \begin{bmatrix} A_2 w_2^r \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} b \\ 1 \end{bmatrix}
\]

Certificate of optimality: given a basic feasible solution, all variables have non-negative reduced costs
Consider a linear program in standard form

\[
\min c^T x \\
\text{s.t. } Ax = b, (\pi) \\
x \geq 0
\]

Given a basis \( B \), when is it optimal?

1. \( B^{-1} b \geq 0 \)
2. \( c_B^T - \pi^T A \geq 0 \)

where \( c_B \) correspond to coefficients of basic variables
Reduced Costs

Given a basic feasible solution, criterion for new variable to enter is negative reduced cost

- Reduced cost of $\lambda^j_1$:

$$c_1^T x^j_1 - \begin{bmatrix} \pi^T & t_1 & t_2 \end{bmatrix} \begin{bmatrix} A_1 x^j_1 \\ 1 \\ 0 \end{bmatrix} = (c_1^T - \pi^T A_1) x^j_1 - t_1$$

- Reduced cost of $\mu^r_1$:

$$c_1^T w^r_1 - \begin{bmatrix} \pi^T & t_1 & t_2 \end{bmatrix} \begin{bmatrix} A_1 x^j_1 \\ 0 \\ 0 \end{bmatrix} = (c_1^T - \pi^T A_1) x^j_1$$

- Similarly for $\lambda^j_2, \mu^r_2$
Idea of the Algorithm: Subproblems

Idea: instead of looking at reduced cost of every variable $\lambda_1^j$, $\lambda_2^j$, $\mu_1^r$, $\mu_2^r$ (there is an enormous number) we can solve the following problems

\[
\begin{align*}
z_1 &= \min \left( c_1^T - \pi^T A_1 \right) x_1 \\
\text{s.t. } &B_1 x_1 = d_1 \\
x_1 &\geq 0
\end{align*}
\]

\[
\begin{align*}
z_2 &= \min \left( c_2^T - \pi^T A_2 \right) x_2 \\
\text{s.t. } &B_2 x_2 = d_2 \\
x_2 &\geq 0
\end{align*}
\]
Three Possibilities

Given the solution of subproblem 1

1. Optimal cost is $-\infty$
   - Simplex output: extreme ray $w'_1$ with $(c_1^T - \pi^T A_1)w'_1 < 0$
   - Conclusion: reduced cost of $\mu'_1$ is negative
   - Action: include $\mu'_1$ in master problem with column
     \[
     \begin{bmatrix}
     A_1 w'_1 \\
     0 \\
     0
     \end{bmatrix}
     \]

2. Optimal cost finite, less then $t_1$
   - Simplex output: extreme point $x'_1$ with $(c_1^T - \pi^T A_1)x'_1 < t_1$
   - Conclusion: reduced cost of $\lambda'_1$ is negative
   - Action: include $\lambda'_1$ in master problem with column
     \[
     \begin{bmatrix}
     A_1 x'_1 \\
     1 \\
     0
     \end{bmatrix}
     \]
Optimal cost is finite, no less than $t_1$

- Conclusion: $(c_1^T - \pi^T A_1)x^j_1 \geq t_1$ for all extreme points $x^j_1$,
  $(c_1^T - \pi^T A_1)w^r_1 \geq 0$ for all extreme rays $w^r_1$

- Action: terminate, we have an optimal basis

Same idea applies to subproblem 2
Idea of the Algorithm: Master

Idea: instead of solving **full master** for all variables, solve **restricted master problem** for ‘worthwhile’ subset of variables

\[ \tilde{J}_1 \subset J_1, \tilde{J}_2 \subset J_2, \tilde{R}_1 \subset R_1, \tilde{R}_2 \subset R_2 \]

\[
\begin{align*}
    z &= \min \sum_{j \in \tilde{J}_1} \lambda^j_1 c_1^T x_1^j + \sum_{r \in \tilde{R}_1} \mu^r_1 c_1^T w_1^r + \sum_{j \in \tilde{J}_2} \lambda^j_2 c_2^T x_2^j + \sum_{r \in \tilde{R}_2} \mu^r_2 c_2^T w_2^r \\
    \sum_{j \in \tilde{J}_1} \lambda^j_1 A_1 x_1^j + \sum_{r \in \tilde{R}_1} \mu^r_1 A_1 w_1^r + \sum_{j \in \tilde{J}_2} \lambda^j_2 A_2 x_2^j + \sum_{r \in \tilde{R}_2} \mu^r_2 A_2 w_2^r &= b \\
    \sum_{j \in \tilde{J}_1} \lambda^j_1 &= 1, \sum_{j \in \tilde{J}_2} \lambda^j_2 = 1 \\
    \lambda^j_1, \lambda^j_2, \mu^r_1, \mu^r_2 &\geq 0
\end{align*}
\]
Dantzig-Wolfe Decomposition Algorithm

1. Solve restricted master with initial basic feasible solution, store $\pi, t_1, t_2$

2. Solve subproblems 1 and 2. If $(c_1^T - \pi^T A_1)x \geq t_1$ and $(c_2^T - \pi^T A_2)x \geq t_2$ terminate with optimal solution:

$$x_1 = \sum_{j \in \tilde{J}_1} \lambda_j^1 x_j^1 + \sum_{r \in \tilde{R}_1} \mu_r^1 w_r^1$$

$$x_2 = \sum_{j \in \tilde{J}_2} \lambda_j^2 x_j^2 + \sum_{r \in \tilde{R}_2} \mu_r^2 w_r^2$$

3. If subproblem $i$ is unbounded, add $\mu_r^i$ to the master

4. If subproblem $i$ has bounded optimal cost less than $t_i$, add $\lambda_j^i$ to the master

5. Generate column associated with entering variable, solve master, store $\pi, t_1, t_2$ and go to step 2
Applicability of the Method

Analysis generalizes to multiple subproblems:

\[
\begin{align*}
\min & \quad c_1^T x_1 + c_2^T x_2 + \cdots + c_t^T x_K \\
\text{s.t.} & \quad A_1 x_1 + A_2 x_2 + \cdots + A_t x_K = b \\
& \quad B_i x_i = d_i, \; i = 1, \ldots, K \\
& \quad x_1, x_2, \ldots, x_K \geq 0
\end{align*}
\]

Approach applies for \( K = 1 \), apply when \textit{subproblem has special structure}

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad Ax = b \\
& \quad Bx = d \\
& \quad x \geq 0
\end{align*}
\]
Dantzig-Wolfe Bounds

Denote:

- $z_i$: optimal objective function value of subproblem $i$, $i = 1, \ldots, K$
- $z^*$: optimal objective function value of problem
- $z$: optimal objective function value of restricted master
- $t_i$: dual optimal multiplier of $\sum_{j \in \tilde{J}_i} \lambda^j_i = 1$ in restricted master

We get bounds at each iteration

- Upper bound:
  \[ z \geq z^* \]

- Lower bound:
  \[ z + \sum_{i=1}^{K} (z_i - t_i) \leq z^* \]
The solution of the restricted master problem is a feasible solution to the original problem.
Proof of Lower Bound \((K = 2)\)

Consider the dual of the master problem:

\[
\begin{align*}
\max & \quad \pi^T b + t_1 + t_2 \\
\text{s.t.} & \quad \pi^T A_1 x_1^j + t_1 \leq c_1^T x_1^j, j \in J_1, (\lambda_1^j) \\
& \quad \pi^T A_1 w_r^1 \leq c_1^T w_r^1, r \in R_1, (\mu_1^r) \\
& \quad \pi^T A_2 x_2^j + t_2 \leq c_2^T x_2^j, j \in J_2, (\lambda_2^j) \\
& \quad \pi^T A_2 w_r^2 \leq c_2^T w_r^2, r \in R_2, (\mu_2^r)
\end{align*}
\]
Note that if $z_1$ is finite

$$z_1 \leq c^T_1 x^j_1 - \pi^T A_1 x^j_1, \forall j \in J_1$$

$$c^T_1 w^r_1 - \pi^T A_1 w^r_1 \geq 0, \forall r \in R_1$$

Same observation holds true for $z_2$ finite

Conclusion: $(\pi, z_1, z_2)$ is feasible for above problem

Weak duality:

$$z^* \geq \pi^T b + z_1 + z_2 = z + (z_1 - t_1) + (z_2 - t_2)$$
1. Algorithm Description [Infanger, Bertsimas]

2. Examples [Bertsimas]

3. Application of Dantzig-Wolfe in Stochastic Programming [BL, §5.5]
   - Reformulation of 2-Stage Stochastic Program
   - Algorithm Description

   - Dantzig-Wolfe Reformulation
   - Relationship to Lagrange Relaxation
Example 1

\[
\begin{align*}
\text{min} & \quad -4x_1 - x_2 - 6x_3 \\
\text{s.t.} & \quad 3x_1 + 2x_2 + 4x_3 = 17 \\
& \quad 1 \leq x_1 \leq 2 \\
& \quad 1 \leq x_2 \leq 2 \\
& \quad 1 \leq x_3 \leq 2
\end{align*}
\]

Divide constraints as follows:

- Represent \( P = \{ x \in \mathbb{R}^3 | 1 \leq x_i \leq 2 \} \) by its extreme points \( x^j \)
- Complicating constraints \( Ax = b, A = \begin{bmatrix} 3 & 2 & 4 \end{bmatrix}, b = 17 \)
Initialization: pick extreme points $x^1 = (2, 2, 2)$, $x^2 = (1, 1, 2)$ with restricted master problem basic variables $\lambda^1, \lambda^2$

Basis matrix:

$$B = \begin{bmatrix} 3 \cdot 2 + 2 \cdot 2 + 4 \cdot 2 & 3 \cdot 1 + 2 \cdot 1 + 4 \cdot 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 18 & 13 \\ 1 & 1 \end{bmatrix}$$

Restricted master:

$$\min -22\lambda^1 - 17\lambda^2$$

s.t. $18\lambda^1 + 13\lambda^2 = 17, (\pi)$

$\lambda^1 + \lambda^2 = 1, (t)$

$\lambda^1, \lambda^2 \geq 0$

Optimal solution $\lambda^1 = 0.8, \lambda^2 = 0.2$, optimal multipliers:

$$\pi = -1, t = -4$$
Objective function coefficients: \( c^T - \pi^T A = \begin{bmatrix} -4 & -1 & -6 \end{bmatrix} - (-1) \begin{bmatrix} 3 & 2 & 4 \end{bmatrix} = \begin{bmatrix} -1 & 1 & -2 \end{bmatrix} \)

Subproblem:

\[
\begin{align*}
\min -x_1 + x_2 - 2x_3 \\
\text{s.t. } 1 \leq x_1 \leq 2, \ 1 \leq x_2 \leq 2, \ 1 \leq x_3 \leq 2
\end{align*}
\]

Optimal solution: \( x^3 = (2, 1, 2) \), objective function value -5 is less than \( t = -4 \)

Introduction of \( \lambda^3 \) to master with coefficients

\[
\begin{bmatrix} 3 \cdot 2 + 2 \cdot 1 + 4 \cdot 2 \\
1 \end{bmatrix} = \begin{bmatrix} 16 \\
1 \end{bmatrix}
\]
Second Iteration: Master

- Restricted master problem:

\[
\begin{align*}
\text{min} & \quad -22\lambda^1 - 17\lambda^2 - 21\lambda^3 \\
\text{s.t.} & \quad 18\lambda^1 + 13\lambda^2 + 16\lambda^3 = 17, \quad (\pi) \\
& \quad \lambda^1 + \lambda^2 + \lambda^3 = 1, \quad (t) \\
& \quad \lambda^1, \lambda^2, \lambda^3 \geq 0
\end{align*}
\]

- Optimal solution \(\lambda^1 = 0.5, \lambda^3 = 0.5\), optimal multipliers:

\(\pi = -0.5, \quad t = -13\)
Second Iteration: Subproblem

- Subproblem:

  \[
  \begin{align*}
  \text{min} & \quad -2.5x_1 - 4x_3 \\
  \text{s.t.} & \quad 1 \leq x_1 \leq 2, \quad 1 \leq x_2 \leq 2, \quad 1 \leq x_3 \leq 2
  \end{align*}
  \]

- Optimal solution: \( x^1 = (2, 2, 2) \), objective function value -13 is equal to \( t = -13 \)

- Optimal solution is

  \[
  x = \frac{1}{2}x^1 + \frac{1}{2}x^3 = \begin{bmatrix} 2 \\ 1.5 \\ 2 \end{bmatrix}
  \]
Graphical Illustration of Example 1

\[ x^2 = (1, 1, 2) \quad (1, 2, 2) \]
\[ x^3 = (2, 1, 2) \]
\[ x^1 = (2, 2, 2) \]

\[ (1, 1, 1) \quad (1, 2, 1) \quad (2, 1, 1) \quad (2, 2, 1) \]
Cube is $P$

Shaded triangle is intersection of $P$ with $3x_1 + 2x_2 + 4x_3 = 17$

Point A: result of first basis ($\lambda^1 = 0.8$, $\lambda^2 = 0.2$)

$x^3$: extreme point brought into master after completion of first iteration

Point B: result of second basis ($\lambda^1 = 0.5$, $\lambda^3 = 0.5$)
Recall solutions at first iteration:

- $z = -21$
- $t = -4$
- $z_1 = -5$

Bounds:

$$-21 \geq z^* \geq -21 + (-5) - (-4) = -22$$

Indeed, $z^* = -21.5$
Example 2

min $-5x_1 + x_2$

s.t. $x_1 \leq 8$

$x_1 - x_2 \leq 4$

$2x_1 - x_2 \leq 10$

$x_1, x_2 \geq 0$

Introduce slack variable $x_3$:

min $-5x_1 + x_2$

s.t. $x_1 + x_3 = 8$

$x_1 - x_2 \leq 4$

$2x_1 - x_2 \leq 10$

$x_1, x_2, x_3 \geq 0$
Decomposition of Example 2

- Treat $x_1 + x_3 = 8$ as a coupling constraint
- $P_1 = \{(x_1, x_2) | x_1 - x_2 \leq 4, 2x_1 - x_2 \leq 10, x_1, x_2 \geq 0\}$
  - Extreme points: $x^1_1 = (6, 2), x^2_1 = (4, 0), x^3_1 = (0, 0)$
  - Extreme rays: $w^1_1 = (1, 2), w^2_1 = (0, 1)$
- $P_2 = \{x_3 | x_3 \geq 0\}$
  - Unique extreme ray: $w^1_2 = 1$
First Iteration: Master

- Initialization: pick extreme point $x_1^1 = (6, 2)$, extreme ray $w_2^1 = 1$

- Restricted master:

$$
\begin{align*}
\min & -28\lambda_1^1 \\
\text{s.t.} & 6\lambda_1^1 + \mu_2^1 = 8, (\pi) \\
\lambda_1^1 = & 1, (t_1) \\
\lambda_1^1, \mu_2^1 \geq & 0
\end{align*}
$$

- Optimal solution $\lambda_1^1 = 1, \mu_2^1 = 2$, optimal multipliers: $\pi = 0, t_1 = -28$
Objective function coefficients:

\[ c_1^T - \pi^T A_1 = \begin{bmatrix} -5 & 1 \end{bmatrix} - (0) \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} -5 & 1 \end{bmatrix} \]

Subproblem:

\[
\begin{align*}
\text{min} & \quad -5x_1 + x_2 \\
\text{s.t.} & \quad x_1 - x_2 \leq 4, \quad 2x_1 - x_2 \leq 10 \\
& \quad x_1, x_2 \geq 0
\end{align*}
\]

Optimal solution: \( w_1^1 = (1, 2) \), objective function value \(-\infty\)
Second Iteration: Master

- **Restricted master problem:**

\[
\begin{align*}
\text{min} & \quad -28\lambda_1^1 - 3\mu_1^1 \\
\text{s.t.} & \quad 6\lambda_1^1 + \mu_1^1 + \mu_2^1 = 8, (\pi) \\
& \quad \lambda_1^1 = 1, (t_1) \\
& \quad \lambda_1^1, \mu_1^1, \mu_2^1 \geq 0
\end{align*}
\]

- **Optimal solution** \( \lambda_1^1 = 1, \mu_1^1 = 2, \mu_2^1 = 0 \), optimal multipliers: \( \pi = -3, t_1 = -10 \)
Second Iteration: Subproblems

- Subproblem:

  \[
  \begin{align*}
  & \text{min } -2x_1 + x_2 \\
  & \text{s.t. } x_1 - x_2 \leq 4, 2x_1 - x_2 \leq 10 \\
  & x_1, x_2 \geq 0
  \end{align*}
  \]

- Optimal solution: \( x = (8, 6) \), objective function value -10 is equal to \( z_1 = -10 \)

- Reduced cost of \( \mu_2^1 \) is 3 (non-negative)

- Optimal solution is

  \[
  x_1^1 + 2w_1^1 = \begin{bmatrix} 8 \\ 6 \end{bmatrix}
  \]
Graphical Illustration of Example 2

- $x_1^1, x_1^2, x_1^3$: extreme points of $P_1$
- $w_1^1, w_1^2$: extreme rays of $P_1$
- Algorithm starts at $(x_1, x_2) = (6, 2)$, reaches optimal solution $(x_1, x_2) = (8, 6)$ after one iteration
Table of Contents

1. Algorithm Description [Infanger, Bertsimas]

2. Examples [Bertsimas]

3. Application of Dantzig-Wolfe in Stochastic Programming [BL, §5.5]
   - Reformulation of 2-Stage Stochastic Program
   - Algorithm Description

   - Dantzig-Wolfe Reformulation
   - Relationship to Lagrange Relaxation
Extended Form 2-Stage Stochastic Program

Primal problem: appropriate for L-shaped method

\[
\min c^T x + \sum_{k=1}^{K} p_k q_k^T y_k \\
\text{s.t. } Ax = b, (\rho) \\
T_k x + Wy_k = h_k, (\pi_k) \\
x, y_k \geq 0
\]

Dual problem: appropriate for Dantzig-Wolfe decomposition

\[
\max \rho^T b + \sum_{k=1}^{K} \pi_k^T h_k \\
\text{s.t. } \rho^T A + \sum_{k=1}^{K} \pi_k^T T_k \leq c^T, (x) \\
\pi_k^T W \leq p_k q_k^T, (y_k)
\]
Consider feasible region of

\[
\begin{bmatrix}
\pi_1^T & \cdots & \pi_K^T
\end{bmatrix}
\begin{bmatrix}
W & \cdots & 0 \\
0 & \cdots & 0 \\
0 & \cdots & W
\end{bmatrix}
\leq
\begin{bmatrix}
q_1^T & \cdots & q_K^T
\end{bmatrix}
\]

Denote $\pi^j, j \in J$ as extreme points, $w^r, r \in R$ as extreme rays

\[
E_j = (\pi^j)^T
\begin{bmatrix}
p_1 T_1 \\
\vdots \\
p_K T_K
\end{bmatrix},
\quad e_j = (\pi^j)^T
\begin{bmatrix}
p_1 h_1 \\
\vdots \\
p_K h_K
\end{bmatrix},
\]

\[
D_r = (w^r)^T
\begin{bmatrix}
p_1 T_1 \\
\vdots \\
p_K T_K
\end{bmatrix},
\quad d_r = (w^r)^T
\begin{bmatrix}
p_1 h_1 \\
\vdots \\
p_K h_K
\end{bmatrix}
\]
Dantzig-Wolfe Full Master Problem

\[ z^* = \max \rho^T b + \sum_{j \in J} \lambda^j e_j + \sum_{r \in R} \mu^r d_r \]

s.t. \( \rho^T A + \sum_{j \in J} \lambda^j E_j + \sum_{r \in R} \mu^r D_r \leq c^T, (x) \)

\[ \sum_{j \in J} \lambda^j = 1, (\theta) \]

\( \lambda^j, \mu^r \geq 0 \)
The dual of the Dantzig-Wolfe full master is

\[
\min c^T x + \theta \\
\text{s.t. } Ax = b \\
E_j x + \theta \geq e_j, j \in J \\
D_r x \geq d_r, r \in R \\
x \geq 0
\]

This is the L-shaped full master problem.
Reduced Costs

We want to bring in

- $\lambda^j$ for which $e_j - E_j x - \theta > 0$
- $\mu^r$ for which $d_r - D_r x > 0$

In order to maximize reduced cost, we need to maximize

$$\sum_{k=1}^{K} (\pi_k)^T h_k - \sum_{k=1}^{K} (\pi_k)^T T_k x$$

where $\pi_k \in \mathbb{R}^{m_k}$
The duals of the Dantzig-Wolfe subproblems are the primal L-shaped subproblems:

\[
\begin{align*}
    z_k &= \max \pi_k^T (h_k - T_k x) \\
    \text{s.t. } \pi_k^T W &\leq q_k, (y_k)
\end{align*}
\]

The duals of the Dantzig-Wolfe subproblems are the primal L-shaped subproblems:

\[
\begin{align*}
    \min q_k^T y_k \\
    \text{s.t. } Wy_k &= h_k - T_k x \\
    y_k &\geq 0
\end{align*}
\]
Master (where $\tilde{J} \subset J$, $\tilde{R} \subset R$)

$$
\begin{align*}
\max z &= \rho^T b + \sum_{j=1}^{\lvert \tilde{J} \rvert} \lambda^j e_j + \sum_{r=1}^{\lvert \tilde{R} \rvert} \mu^r d_r \\
\text{s.t. } \rho^T A + \sum_{j=1}^{\lvert \tilde{J} \rvert} \lambda^j E_j + \sum_{r=1}^{\lvert \tilde{R} \rvert} \mu^r D_r &\leq c^T \\
\sum_{j=1}^{\lvert \tilde{J} \rvert} \lambda^j &= 1, \lambda^j \geq 0, \mu^r \geq 0
\end{align*}
$$

Scenario subproblems:

$$
\begin{align*}
\max \pi^T (h_k - T_k x^v) \\
\text{s.t. } \pi^T W &\leq q^T
\end{align*}
$$
Algorithm

Step 0. $|\tilde{J}| = |\tilde{R}| = v = 0$

Step 1. $v = v + 1$ and solve (3) - (5). Let the solution be $(\rho^v, \lambda^v, \mu^v)$ with dual solution $(x^v, \theta^v)$

Step 2. For $k = 1, \ldots, K$, solve (6) - (7)

- If extreme ray $w^v$ is found, set $d_{|\tilde{R}|+1} = (w^v)^T h_k$, $D_{|\tilde{R}|+1} = (w^v)^T T_k$, $|\tilde{R}| = |\tilde{R}| + 1$ and return to step 1
- If all subproblems are solvable, let

\[
E_{|\tilde{J}|+1} = \sum_{k=1}^{K} p_k (\pi^v_k)^T T_k, \quad e_{|\tilde{J}|+1} = \sum_{k=1}^{K} p_k (\pi^v_k)^T h_k
\]

- If $e_{|\tilde{J}|+1} - E_{|\tilde{J}|+1} x^v - \theta \leq 0$, then stop with $(\rho^v, \lambda^v, \mu^v)$ and $(x^v, \theta^v)$ optimal
- If $e_{|\tilde{J}|+1} - E_{|\tilde{J}|+1} x^v - \theta^v > 0$, set $|\tilde{J}| = |\tilde{J}| + 1$ and return to step 1
Dantzig-Wolfe Bounds Revisited

- Lower bound: \( z \leq z^* \)
- Upper bound: \( z^* \leq c^T x + \sum_{k=1}^{K} p_k z_k \)
- Dantzig-Wolfe bounds are the same as the L-shaped bounds
Dantzig-Wolfe Versus L-Shaped Method

- Both algorithms go through the same steps
- Difference: we solve the dual problems instead of the primal problems
# Table of Contents

1. Algorithm Description [Infanger, Bertsimas]

2. Examples [Bertsimas]

3. Application of Dantzig-Wolfe in Stochastic Programming [BL, §5.5]
   - Reformulation of 2-Stage Stochastic Program
   - Algorithm Description

   - Dantzig-Wolfe Reformulation
   - Relationship to Lagrange Relaxation
Integer Programming Formulation

\[(IP) : \min \{ c^T x : x \in X \} \]
\[X = Y \cap Z \]
\[Y = \{ Dx \geq d \} \]
\[Z = \{ Bx \geq b \} \cap \mathbb{Z}^n \]

**Structural assumption:** \( \text{OPT}(Z, c) : \{ \min c^T x : x \in Z \} \) can be solved rapidly in practice
Idea: Apply Dantzig-Wolfe to (IP) using Minkowski Representation Theorem to represent

\[ \text{conv}(Z) = \text{conv}(\{ Bx \geq b \} \cap \mathbb{Z}^n) \]

\( \text{conv}(Z) \) is the gray area
Dantzig-Wolfe Reformulation

\[(DWC) : z^{DWc} = \min_{\lambda \geq 0} \sum_{j \in J} (c^T x^j) \lambda^j\]

s.t. \[\sum_{j \in J} (Dx^j) \lambda^j \geq d\]

\[\sum_{j \in J} \lambda^j = 1, \quad \sum_{j \in J} x^j \lambda^j \in \mathbb{Z}^n\]

where

- \(x^j\) is the set of extreme points of \(\text{conv}(Z)\),
- \(\text{conv}(Z)\) is the convex hull of \(Z\)
- \(J\) is the set of extreme points of \(\text{conv}(Z)\)
The linear relaxation of \((DWc)\) is called the **Master Linear Program (MLP)**

When we only consider a subset \(\tilde{J} \subset J\) of the extreme points of \(\text{conv}(Z)\) we get the **Restricted Master Linear Program (RMLP)**

\[
(RMLP) : z^{RMLP} = \min_{\lambda \geq 0} \sum_{j \in \tilde{J}} (c^T x^j) \lambda^j \\
\text{s.t. } \sum_{j \in \tilde{J}} (Dx^j) \lambda^j \geq d, (\pi) \\
\sum_{j \in \tilde{J}} \lambda^j = 1, (\sigma)
\]
Observations

1. The reduced cost associated to $\lambda^j$ is $c^T x^j - \pi^T Dx^j - \sigma$

2. **Important:** $z = \min_{j \in \tilde{J}} (c^T x^j - \pi^T Dx^j) = \min_{x \in Z} (c^T - \pi^T D)x = \min_{Bx \geq b, x \in \mathbb{Z}^n_+} (c^T - \pi^T D)x$ is an easy integer program

3. $z^{RMLP} = \sum_{j \in \tilde{J}} (c^T x^j)\lambda^j$ is an upper bound on $z_{MLP}$ and (MLP) is solved when $z - \sigma = 0$

4. If solution $\lambda$ of (RMLP) is integer, $z^{RMLP}$ is an upper bound for (IP)
Column Generation Algorithm for (MLP)

1. Initialize primal and dual bounds $UB = +\infty$, $LB = -\infty$

2. Iteration $t$
   - Solve $(RMLP)$ over $x^j, j \in \tilde{J}^t$, record primal solution $\lambda^t$ and dual solution $(\pi^t, \sigma^t)$
   - Solve pricing problem $(SP^t)$: $z^t = \min \{(c^T - (\pi^t)^T D)x : x \in Z\}$, let $x^t$ be an optimal solution. If $z^t - \sigma^t = 0$ set $UB = z^{RMLP}$ and stop with optimal solution to $(MLP)$. Else, add $x^t$ to $\tilde{J}^t$ in $(RMLP)$.
   - Compute lower bound $(\pi^t)^T d + z^t$. Update $LB = \max\{LB, (\pi^t)^T d + z^t\}$. If $LB = UB$, stop with optimal solution to $(MLP)$

3. Increment $t$, return to step 2
Relationship to Lagrange Relaxation

Relaxing ‘difficult’ constraints $Dx \geq d$, while keeping the remaining constraints $Z = \{ x \in \mathbb{Z}^n_+ : Bx \geq b \}$, we get

- the dual function

$$g(\pi) = \min_{x} \{ c^T x + \pi^T (d - Dx) : Bx \geq b, x \in \mathbb{Z}^n_+ \} \quad (8)$$

- the dual bound

$$z_{LD} = \max_{\pi \geq 0} g(\pi) = \max_{\pi \geq 0} \min_{x \in Z} \{ c^T x + \pi^T (d - Dx) \}$$
Reformulation of Dual Bound

\[ z_{LD} = \max_{\pi \geq 0} \min_{j \in J} \{ c^T x^j + \pi^T (d - Dx^j) \} \]

where

- \( x^j \) is the set of extreme points of \( \text{conv}(Z) \),
- \( \text{conv}(Z) \) is the convex hull of \( Z \)
- \( J \) is the set of extreme points of \( \text{conv}(Z) \)

Equivalently:

\[ z_{LD} = \max_{\pi \geq 0, \sigma} \pi^T d + \sigma \]

s.t. \( \sigma \leq c^T x^j - \pi^T Dx^j \), \( j \in J \), \( (\lambda^i) \)
Taking the dual:

\[ z_{LD} = \min_{\lambda^j \geq 0, j \in J} \sum_{j \in J} (c^T x^j) \lambda^j \]  \quad (9)

s.t. \[ \sum_{j \in J} (Dx^j) \lambda^j \geq d, (\pi) \]  \quad (10)

\[ \sum_{j \in J} \lambda^j = 1, (\sigma) \]  \quad (11)
Relationship Between Lagrange Dual Bound and LP Relaxation of Dantzig-Wolfe Reformulation

- **Observe:** The linear program (9) - (11) is the master linear program (MLP) of Dantzig-Wolfe

- **Conclusion:** Solving the Lagrange Relaxation (9) - (11) will give the same bound as solving (MLP) using Dantzig-Wolfe decomposition