15.081J/6.251J Introduction to Mathematical Programming

Lecture 15: Large Scale Optimization, II

1 Outline

SLIDE 1

1. Dantzig-Wolfe decomposition
2. Key Idea
3. Bounds

2 Decomposition

SLIDE 2

$$\begin{aligned} \min \quad & \boldsymbol{c}_1'\boldsymbol{x}_1 + \boldsymbol{c}_2'\boldsymbol{x}_2 \\ \text{s.t.} \quad & \boldsymbol{D}_1\boldsymbol{x}_1 + \boldsymbol{D}_2\boldsymbol{x}_2 = \boldsymbol{b}_0 \\ & \boldsymbol{F}_1\boldsymbol{x}_1 = \boldsymbol{b}_1 \\ & \boldsymbol{F}_2\boldsymbol{x}_2 = \boldsymbol{b}_2 \\ & \boldsymbol{x}_1, \boldsymbol{x}_2 \geq \boldsymbol{0} \end{aligned}$$

- Relation with stochastic programming?
- Firm's problem

2.1 Reformulation

SLIDE 3

- $P_i = \{\boldsymbol{x}_i \geq \boldsymbol{0} \mid \boldsymbol{F}_i\boldsymbol{x}_i = \boldsymbol{b}_i\}$, $i = 1, 2$
- $\boldsymbol{x}_i^j$, $j \in J_i$ extreme points of P_i
- $\boldsymbol{w}_i^k$, $k \in K_i$, extreme rays of P_i.
- For all $\boldsymbol{x}_i \in P_i$

$$\boldsymbol{x}_i = \sum_{j \in J_i} \lambda_i^j \boldsymbol{x}_i^j + \sum_{k \in K_i} \theta_i^k \boldsymbol{w}_i^k,$$

$\lambda_i^j \geq 0$ and $\theta_i^k \geq 0$

$$\sum_{j \in J_i} \lambda_i^j = 1, \qquad i = 1, 2$$

SLIDE 4

$$\begin{aligned} \min \quad & \sum_{j \in J_1} \lambda_1^j \boldsymbol{c}_1'\boldsymbol{x}_1^j + \sum_{k \in K_1} \theta_1^k \boldsymbol{c}_1'\boldsymbol{w}_1^k + \sum_{j \in J_2} \lambda_2^j \boldsymbol{c}_2'\boldsymbol{x}_2^j + \sum_{k \in K_2} \theta_2^k \boldsymbol{c}_2'\boldsymbol{w}_2^k \\ \text{s.t.} \quad & \sum_{j \in J_1} \lambda_1^j \boldsymbol{D}_1\boldsymbol{x}_1^j + \sum_{k \in K_1} \theta_1^k \boldsymbol{D}_1\boldsymbol{w}_1^k + \sum_{j \in J_2} \lambda_2^j \boldsymbol{D}_2\boldsymbol{x}_2^j \\ & \qquad + \sum_{k \in K_2} \theta_2^k \boldsymbol{D}_2\boldsymbol{w}_2^k = \boldsymbol{b}_0 \\ & \sum_{j \in J_1} \lambda_1^j = 1 \\ & \sum_{j \in J_2} \lambda_2^j = 1 \\ & \lambda_i^j \geq 0,\ \theta_i^k \geq 0, \qquad \forall\ i, j, k. \end{aligned}$$

Huge # variables, $m_0 + 2$ constraints SLIDE 5

- A bfs is available with a basis matrix $\boldsymbol{B}$
- $\boldsymbol{p}' = \boldsymbol{c}_B'\boldsymbol{B}^{-1}$; $\boldsymbol{p} = (\boldsymbol{q}, r_1, r_2)$
- Is $\boldsymbol{B}$ optimal?
- Check reduced costs

$$(\boldsymbol{c}_1' - \boldsymbol{q}'\boldsymbol{D}_1)\boldsymbol{x}_1^j - r_1$$
$$(\boldsymbol{c}_1' - \boldsymbol{q}'\boldsymbol{D}_1)\boldsymbol{w}_1^k$$

- Huge number of them

3 Key idea

SLIDE 6

Consider subproblem:

$$\begin{array}{ll} \min & (\boldsymbol{c}_1' - \boldsymbol{q}'\boldsymbol{D}_1)\boldsymbol{x}_1 \\ \text{s.t.} & \boldsymbol{x}_1 \in P_1, \end{array}$$

- If optimal cost of subproblem is $-\infty$, an extreme ray $\boldsymbol{w}_1^k$ is generated: $(\boldsymbol{c}_1' - \boldsymbol{q}'\boldsymbol{D}_1)\boldsymbol{w}_1^k < 0$, i.e., reduced cost of θ_1^k is negative; Generate column $[\boldsymbol{D}_1\boldsymbol{w}_1^k, 0, 0]'$
- If optimal cost is finite and smaller than r_1, then, an extreme point $\boldsymbol{x}_1^j$ is generated: $(\boldsymbol{c}_1' - \boldsymbol{q}'\boldsymbol{D}_1)\boldsymbol{x}_1^j < r_1$, i.e., reduced cost of λ_1^j is negative; Generate column $[\boldsymbol{D}_1\boldsymbol{x}_1^j, 1, 0]'$
- Otherwise, reduced costs are nonnegative
- Repear for subproblem:

$$\begin{array}{ll} \min & (\boldsymbol{c}_2' - \boldsymbol{q}'\boldsymbol{D}_2)\boldsymbol{x}_2 \\ \text{s.t.} & \boldsymbol{x}_2 \in P_2, \end{array}$$

4 Remarks

SLIDE 7

- Economic interpretation
- Applicability of the method

$$\begin{array}{lll} \min & \boldsymbol{c}_1'\boldsymbol{x}_1 + \boldsymbol{c}_2'\boldsymbol{x}_2 + \cdots + \boldsymbol{c}_t'\boldsymbol{x}_t & \\ \text{s.t.} & \boldsymbol{D}_1\boldsymbol{x}_1 + \boldsymbol{D}_2\boldsymbol{x}_2 + \cdots + \boldsymbol{D}_t\boldsymbol{x}_t = \boldsymbol{b}_0 & \\ & \boldsymbol{F}_i\boldsymbol{x}_i = \boldsymbol{b}_i, & i = 1, 2, \ldots, t \\ & \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_t \geq \boldsymbol{0}. & \end{array}$$

$$
\begin{aligned}
\min \quad & \boldsymbol{c}'\boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{D}\boldsymbol{x} = \boldsymbol{b}_0 \\
& \boldsymbol{F}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \geq \boldsymbol{0},
\end{aligned}
$$

4.1 Termination

SLIDE 8

- Finite termination
- Algorithm makes substantial progress in the beginning, but very slow later on
- no faster than the revised simplex method applied to the original problem
- Storage with t subproblems
- Original: $O\big((m_0 + tm_1)^2\big)$
- Decomposition algorithm $O\big((m_0+t)^2\big)$ for the tableau of the master problem, and t times $O(m_1^2)$ for subproblems.
- If $t = 10$ and if $m_0 = m_1$ is much larger than t, memory requirements for decomposition algorithm are about 100 times smaller than revised simplex method.

5 Example

SLIDE 9

-
$$
\begin{aligned}
\min \quad & -4x_1 - x_2 - 6x_3 \\
\text{s.t.} \quad & 3x_1 + 2x_2 + 4x_3 = 17 \\
& 1 \leq x_1 \leq 2 \\
& 1 \leq x_2 \leq 2 \\
& 1 \leq x_3 \leq 2.
\end{aligned}
$$
- $P = \{\boldsymbol{x} \in \Re^3 \mid 1 \leq x_i \leq 2,\ i = 1, 2, 3\}$; eight extreme points;
- Master problem:

$$
\sum_{j=1}^{8} \lambda^j \boldsymbol{D}\boldsymbol{x}^j = 17,
$$

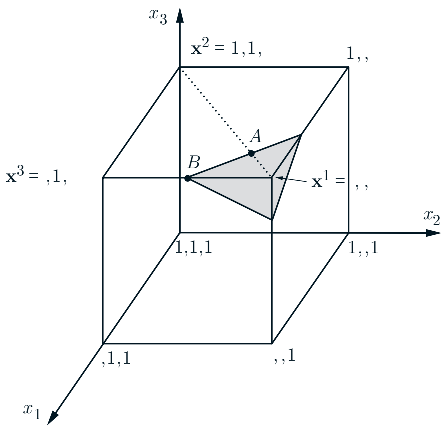
$$
\sum_{j=1}^{8} \lambda^j = 1,
$$

SLIDE 10

- $\boldsymbol{x}^1 = (2, 2, 2)$ and $\boldsymbol{x}^2 = (1, 1, 2)$; $\boldsymbol{D}\boldsymbol{x}^1 = 18$, $\boldsymbol{D}\boldsymbol{x}^2 = 13$

- $\boldsymbol{B} = \begin{bmatrix} 18 & 13 \\ 1 & 1 \end{bmatrix}$; $\boldsymbol{B}^{-1} = \begin{bmatrix} 0.2 & -2.6 \\ -0.2 & 3.6 \end{bmatrix}$
-
$$c_{B(1)} = \boldsymbol{c}'\boldsymbol{x}^1 = \begin{bmatrix} -4 & -1 & -6 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = -22,$$
$$c_{B(2)} = \boldsymbol{c}'\boldsymbol{x}^2 = \begin{bmatrix} -4 & -1 & -6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = -17.$$
- $\boldsymbol{p}' = \begin{bmatrix} \boldsymbol{q}' & r \end{bmatrix} = \boldsymbol{c}'_B \boldsymbol{B}^{-1} = \begin{bmatrix} -22 & -17 \end{bmatrix} \boldsymbol{B}^{-1} = \begin{bmatrix} -1 & -4 \end{bmatrix}$.
-
$$\boldsymbol{c}' - \boldsymbol{q}'\boldsymbol{D} = \begin{bmatrix} -4 & -1 & -6 \end{bmatrix} - (-1)\begin{bmatrix} 3 & 2 & 4 \end{bmatrix} = [-1 \;\; 1 \;\; -2],$$
optimal solution is $\boldsymbol{x}^3 = (2, 1, 2)$ with optimal cost $-5 \leq r = -4$
- Generate the column corresponding to λ^3.

SLIDE 11



6 Starting the algorithm

SLIDE 12

$$\begin{aligned} \min \quad & \sum_{t=1}^{m_0} y_t \\ \text{s.t.} \quad & \sum_{i=1,2} \left(\sum_{j \in J_i} \lambda_i^j \boldsymbol{D}_i \boldsymbol{x}_i^j + \sum_{k \in K_i} \theta_i^k \boldsymbol{D}_i \boldsymbol{w}_i^k \right) + \boldsymbol{y} = \boldsymbol{b}_0 \\ & \sum_{j \in J_1} \lambda_1^j = 1 \end{aligned}$$

$$\sum_{j \in J_2} \lambda_2^j = 1$$

$$\lambda_i^j \geq 0,\ \theta_i^k \geq 0,\ y_t \geq 0, \qquad \forall\, i, j, k, t.$$

7 Bounds

SLIDE 13

- Optimal cost z^*
- z cost of feasible solution obtained at some intermediate stage ofe decomposition algorithm.
- r_i be the value of the dual variable associated with the convexity constraint for the ith subproblem
- z_i optimal cost in the ith subproblem
- Then,

$$z + \sum_i (z_i - r_i) \leq z^* \leq z.$$

7.1 Proof

SLIDE 14

Dual of master problem

$$\begin{array}{rlllll} \max & \boldsymbol{q}'\boldsymbol{b}_0 + r_1 + r_2 & & & & \\ \text{s.t.} & \boldsymbol{q}'\boldsymbol{D}_1\boldsymbol{x}_1^j & + \ r_1 & \leq & \boldsymbol{c}_1'\boldsymbol{x}_1^j, & \forall\, j \in J_1, \\ & \boldsymbol{q}'\boldsymbol{D}_1\boldsymbol{w}_1^k & & \leq & \boldsymbol{c}_1'\boldsymbol{w}_1^k, & \forall\, k \in K_1, \\ & \boldsymbol{q}'\boldsymbol{D}_2\boldsymbol{x}_2^j & + \ r_2 & \leq & \boldsymbol{c}_2'\boldsymbol{x}_2^j, & \forall\, j \in J_2, \\ & \boldsymbol{q}'\boldsymbol{D}_2\boldsymbol{w}_2^k & & \leq & \boldsymbol{c}_2'\boldsymbol{w}_2^k, & \forall\, k \in K_2. \end{array}$$

SLIDE 15

- $(\boldsymbol{q}, r_1, r_2)$ dual variables

$$\boldsymbol{q}'\boldsymbol{b}_0 + r_1 + r_2 = z$$

- z_1 is the optimal cost in the first subproblem:

$$\begin{aligned} \min_{j \in J_1} (\boldsymbol{c}_1'\boldsymbol{x}_1^j - \boldsymbol{q}'\boldsymbol{D}_1\boldsymbol{x}_1^j) &= z_1, \\ \min_{k \in K_1} (\boldsymbol{c}_1'\boldsymbol{w}_1^k - \boldsymbol{q}'\boldsymbol{D}_1\boldsymbol{w}_1^k) &\geq 0. \end{aligned}$$

- $(\boldsymbol{q}, z_1, z_2)$ is a feasible solution to the dual of master problem

- By weak duality,

$$\begin{aligned} z^* &\geq \boldsymbol{q}'\boldsymbol{b}_0 + z_1 + z_2 \\ &= \boldsymbol{q}'\boldsymbol{b}_0 + r_1 + r_2 + (z_1 - r_1) + (z_2 - r_2) \\ &= z + (z_1 - r_1) + (z_2 - r_2), \end{aligned}$$

7.2 Example

Slide 16

- $(\lambda^1, \lambda^2) = (0.8, 0.2)$
- $\boldsymbol{c}_B = (-22, -17)$, $z = (-22, -17)'(0.8, 0.2) = -21$
- $r = -4$; $z_1 = (-1, 1, -2)'(2, 1, 2) = -5$.
- $-21 \geq z^* \geq -21 + (-5) - (-4) = -22$
- $z^* = -21.5$