6.215/6.255J/15.093J/IDS.200J Optimization Methods

Lecture 8: Large Scale Optimization

October 5, 2021

Today's Lecture

Outline

- Large scale optimization:
 - Column generation methods
 - Example: The cutting stock problem
 - Cutting plane methods
- Multi-stage optimization

Large-Scale Problems

Consider a linear optimization problem

$$\begin{array}{ll}
\min & \boldsymbol{c}^{\mathsf{T}} \boldsymbol{x} \\
\text{s.t.} & \boldsymbol{A} \boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \ge 0
\end{array}$$

- When is such a problem "large scale"?
 - the number of variables n is too large to represent explicitly; and/or
 - \bullet the number of constraints m is too large to store in system memory.
- What to do?

Large-Scale Problems

Large number of variables. Column generation

Consider the linear optimization problem with n large as the master problem

$$\begin{array}{ll}
\min & \boldsymbol{c}^{\mathsf{T}} \boldsymbol{x} \\
\text{s.t.} & \boldsymbol{A} \boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \ge 0
\end{array}$$

Column generation idea:

- consider I a subset of $1, \ldots, n$.
- find optimal basis **B** (with simplex method) for restricted problem

min
$$\sum_{i \in I} c_i x_i$$

s.t. $\sum_{i \in I} \mathbf{A}_i x_i = \mathbf{b}$
 $\mathbf{x} > 0$

- find a new column A_j , $j \notin I$ with negative reduced cost $\overline{c}_j < 0$. If not possible B is optimal in the master problem.
- update $I := I \cup \{j\}$.

Large-Scale Problems

Meta algorithm: Column generation

A column generation iteration has two main computational tasks

- solving the restricted problems
- dual feasibility problem: verifying dual feasibility; finding a negative reduced cost
 - task 1: done using primal simplex method since basis **B** remains primal feasible in the new restricted problem

min
$$\boldsymbol{c}^T \boldsymbol{x}$$

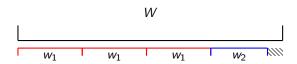
s.t. $\sum_{i \in I} \boldsymbol{A}_i x_i + \boldsymbol{A}_j x_j = \boldsymbol{b},$
 $\boldsymbol{x} > 0.$

• task 2: more critical, the master problem needs to possess some structure which enables us to find negative reduced cost efficiently (without enumerating all columns).

5 / 33

- Company has a supply of large rolls of paper, each of width W.
- b_i rolls of width w_i , i = 1, ..., m need to be produced.
- Example: W=70 can be cut in 3 rolls of width $w_1=17$ and 1 roll of width $w_2=15$, with a waste of

$$70 - (3 \times 17 + 1 \times 15) = 4$$



• Given w_1, \ldots, w_m and W there are many cutting patterns:

in our previous example W=70, $w_1=17$, $w_2=15$, here are two patterns: (3,1) and (2,2)

$$3 \times 17 + 1 \times 15 = 66 \le 70$$

 $2 \times 17 + 2 \times 15 = 64 \le 70$

• In general a pattern is defined as m integers (a_1, \ldots, a_m) such that

$$\sum_{i=1}^m a_i w_i \leq W$$

 But ... there can be a huge number of such patterns. Too many to enumerate.

Problem definition

• Input: Given w_i , b_i , i = 1, ..., m (b_i number of rolls of width w_i demanded by customers), and W (width of large rolls available to the company):

• Problem: Find how to minimize the number of large rolls in order to meet the customers demand.

A concrete example

- W = 70, $w_1 = 20$, $w_2 = 11$, $b_1 = 12$, $b_2 = 17$
- There are 15 feasible patterns:

$$\binom{1}{0},\binom{2}{0},\binom{3}{0},\binom{0}{1},\binom{1}{1},\binom{1}{1},\binom{2}{1},\binom{0}{2},\binom{1}{2},\binom{2}{2},\binom{0}{3},\binom{1}{3},\binom{0}{4},\binom{1}{4},\binom{0}{5},\binom{0}{6}$$

• $x_1,\ldots,x_{15}=\#$ of patterns of type $\binom{1}{0},\ldots,\binom{0}{6}$, respectively

min
$$x_1 + \dots + x_{15}$$

s.t. $x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \dots + x_{15} \begin{pmatrix} 0 \\ 6 \end{pmatrix} \ge \begin{pmatrix} 12 \\ 17 \end{pmatrix}$
 $x_1, \dots, x_{15} > 0$

• rounding or other ad hoc method may be necessary to get integer solution.

A concrete example

min
$$x_1 + \dots + x_{15}$$

s.t. $x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \dots + x_{15} \begin{pmatrix} 0 \\ 6 \end{pmatrix} \ge \begin{pmatrix} 12 \\ 17 \end{pmatrix}$
 $x_1, \dots, x_{15} \ge 0$

Example:

$$2 \binom{0}{6} + 1 \binom{0}{5} + 4 \binom{3}{0} = \binom{12}{17} \quad 7 \text{ rolls used}$$

$$4\binom{0}{4} + \binom{0}{1} + 4\binom{3}{0} = \binom{12}{17}$$
 9 rolls used

Formulation

Decision variables:

 $x_j = \text{number of rolls cut by pattern } j \text{ characterized by vector } A_j$.

min
$$\sum_{j=1}^{n} x_{j}$$

s.t. $\sum_{j=1}^{n} \mathbf{A}_{j} \cdot x_{j} \ge \begin{pmatrix} b_{1} \\ \vdots \\ b_{m} \end{pmatrix} = \mathbf{b}$
 $x_{j} \ge 0 \quad \forall j \quad \text{(integer)}$

- Huge number of variables.
- Let us consider the linear optimization problem relaxation (removing the integrality restriction).
- Can we apply column generation, that is generate the patterns A_j on the fly?

4 D > 4 D > 4 E > 4 E > E = 900

Algorithm

Idea: Generate feasible patterns as needed.

Start with initial patterns (m of them):

$$\begin{pmatrix} \left\lfloor \frac{W}{w_1} \right\rfloor \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \left\lfloor \frac{W}{w_2} \right\rfloor \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \left\lfloor \frac{W}{w_3} \right\rfloor \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \left\lfloor \frac{W}{w_m} \right\rfloor \end{pmatrix}$$

Solve:

min
$$x_1 + \cdots + x_m$$

s.t. $x_1 \mathbf{A}_1 + \cdots + x_m \mathbf{A}_m = \mathbf{b}$
 $x_i \ge 0, \ 1 \le i \le m$

3 Compute reduced costs $\overline{c}_j = 1 - \mathbf{p}^T \mathbf{A}_j$ for all patterns j

If $\overline{c}_j \geq 0$ current set of patterns optimal If $\overline{c}_s < 0 \Rightarrow x_s$ needs to enter basis

Question: How to compute $\overline{c}_j = 1 - \mathbf{p}^T \mathbf{A}_j$ for all j? (huge number !!)

Algorithm - computing reduced costs $\overline{c}_i = 1 - \boldsymbol{p}^T \boldsymbol{A_j}$ for all j

Key idea: Solve the following problem (integer knapsack) problem:

$$z^* = \max \sum_{\substack{i=1 \ m \ a_i}}^m p_i a_i$$

s.t. $\sum_{\substack{i=1 \ a_i \geq 0, \text{ integer}}}^m w_i a_i \leq W$

- If $z^* \le 1 \implies 1 \mathbf{p}^T \mathbf{A}_i > 0 \ \forall j \implies \text{current solution optimal}$
- If $z^* > 1$ \Rightarrow $\exists s: 1 \boldsymbol{p^T} \boldsymbol{A_s} < 0$ \Rightarrow Variable x_s becomes basic, i.e., a new pattern $\boldsymbol{A_s}$ will enter the basis.
- Perform min-ratio test and update the basis (revised simplex method)

←ロト→団ト→重ト→重・ 夕へ○

Solving the integer knapsack problem

Using dynamic programming:

$$F(u) = \max p_1 a_1 + \dots + p_m a_m$$

s.t. $w_1 a_1 + \dots + w_m a_m \le u$
 $a_i \ge 0$, integer

- For $u \leq w_{min}$, F(u) = 0.
- For $u \geq w_{min}$

$$F(u) = \max_{i=1,...,m} \{p_i + F(u - w_i)\}$$

why?



Solving the integer knapsack problem - example

$$\begin{array}{ll} \text{max} & 11x_1 + 7x_2 + 5x_3 + x_4 \\ \text{s.t.} & 6x_1 + 4x_2 + 3x_3 + x_4 \leq 25 \\ & x_i \geq 0, \quad x_i \text{ integer} \end{array}$$

```
• F(0) = 0
                                                                      (There is nothing to cut)
  • F(1) = 1
                                                                         (Only w_4 can be cut)
  • F(2) = 1 + F(1) = 2
                                                                         (Only w_4 can be cut)
  • F(3) = \max(5 + F(0), 1 + F(2)) = 5
                                                                                 (Best cut w_3)
  • F(4) = \max(7 + F(0), 5 + F(1), 1 + F(3)) = 7
                                                                                 (Best cut w_2)
  • F(5) = \max(7 + F(1), 5 + F(2), 1 + F(4)) = 8
                                                                               (Cut w_2 or w_4)
  • F(6) = \max(11 + F(0), 7 + F(2), 5 + F(3), 1 + F(5)) = 11
                                                                             (Best cut off w_1)
  • F(7) = \max(11 + F(1), 7 + F(3), 5 + F(4), 1 + F(6)) = 12
                                                                             (All w_i are equal)
  • F(8) = \max(11 + F(2), 7 + F(4), 5 + F(5), 1 + F(7)) = 14
                                                                             (Best cut off w_2)
  • F(9) = \max(11 + F(3), 7 + F(5), 5 + F(6), 1 + F(8)) = 16
                                                                               (Cut w_1 or w_3)
  • F(10) = \max(11 + F(4), 7 + F(6), 5 + F(7), 1 + F(9)) = 18
                                                                               (Cut w_1 or w_2)
  • F(u) = 11 + F(u - 6) = 16 u > 11
\Rightarrow F(25) = 11 + F(19) = 11 + 11 + F(13) = 11 + 11 + 11 + F(7) = 33 + 12 = 45
```

6.255J © 2021 (MIT)

Large Scale

10/5/2021

Cutting Plane Methods

• Consider the dual of a primal in standard form:

- Large n in primal \Rightarrow large number of constraints in dual
- Let I a subset of $\{1, \ldots, n\}$ and solve the relaxed dual problem:

$$\max_{\text{s.t.}} \mathbf{p}^{\mathsf{T}} \mathbf{b}$$
s.t.
$$\mathbf{p}^{\mathsf{T}} \mathbf{A}_{i} \leq c_{i}, i \in I$$

Cutting Plane Methods

• Let I a subset of $\{1, \ldots, n\}$ and solve the relaxed dual problem:

$$\max_{s.t.} \begin{array}{l} \boldsymbol{p}^T \boldsymbol{b} \\ \boldsymbol{p}^T \boldsymbol{A}_i \le c_i, i \in I \end{array}$$

- If optimal solution of relaxed problem p^* satisfies all constraints of the original problem, then it is optimal for the original problem
- If optimal solution of relaxed problem is infeasible for the original problem, bring a violated constraint into *I*
- Method needs:
 - a way to check feasibility
 - a way to identify violated constraints (the "separation problem") (one possibility is to solve $\min_i \{c_i p^{*T} A_i\}$ over all i.)
- Cutting planes for dual = Column generation for primal

Example

Problem:

	Wrenches	Pliers	Cap.
Steel (lbs)	1.5	1.0	27,000
Molding machine (hrs)	1.0	1.0	21,000
Assembly machine (hrs)	0.3	0.5	9,000
Demand limit (tools/day)	15,000	16,000	
Contribution to earnings	\$125	\$100	
(\$/1000 units)		'	

Formulation:

$$\begin{array}{ll} \textit{W} \colon \# \text{ wrenches; } P \ \# \text{ pliers (} \times 1000\text{)} \\ \text{max} & 125W + 100P \\ \text{s.t.} & \textit{W} \le 15 \\ \textit{P} \le 16 \\ & 1.5W + P \le 27 \\ & \textit{W} + P \le 21 \\ & 0.3W + 0.5P \le 9 \\ & \textit{W}, \textit{P} > 0 \end{array}$$

Example

Problem:

Steel (lbs)		
Molding machine (hrs)		
Assembly machine (hrs)		
Demand limit (tools/day)		
Contribution to earnings		
(\$/1000 units)		

Wrenches	Pliers	Cap.
1.5	1.0	27,000
1.0	1.0	21,000
0.3	0.5	9,000
15,000	16,000	
\$125	\$100	

Formulation:

$$W$$
: # wrenches; P # pliers (×1000)
max 125 W + 100 P
s.t. $W \le 15$
 $P \le 16$
 $1.5W + P \le 27$
 $W + P \le 21$
 $0.3W + 0.5P \le 9$
 $W, P > 0$

Random data:

- Assembly capacity: 8000 with probability 0.5 10,000 with probability 0.5
- Contribution from wrenches: 160 with probability 0.5 with probability 0.5

Time dynamics: A two-stage problem

Decisions

- Stage 1: Need to decide steel capacity in the current quarter. Cost \$58/1000lbs.
 Soon after uncertainty (assembly cap. and wrenches contrib.) will be resolved.
- Stage 2: Next quarter, company will decide production quantities.

Possible states in the second stage

State	Cap.	W. contr.	Prob.
1	8,000	160	0.25
2	10,000	160	0.25
3	8,000	90	0.25
4	10,000	90	0.25

Decision variables: S: steel capacity,

 $P_i, W_i : i = 1, ..., 4$ production plan under state i.

Formulation of the two-stage problem

Formulation

$$\begin{array}{lll} \max & -58S + 0.25Z_1 + 0.25Z_2 + 0.25Z_3 + 0.25Z_4 \\ \mathrm{s.t.} & \mathrm{Ass.} \ 1 & 0.3W_1 + 0.5P_1 \leq 8 \\ & \mathrm{Mol.} \ 1 & W_1 + P_1 \leq 21 \\ & \mathrm{Ste.} \ 1 & -S + 1.5W_1 + P_1 \leq 0 \\ & \mathrm{W.d.} \ 1 & W_1 \leq 15 \\ & \mathrm{P.d.} \ 1 & P_1 \leq 16 \\ & \mathrm{Obj.} \ 1 & -Z_1 + 160W_1 + 100P_1 = 0 \end{array}$$

Formulation of the two-stage problem

and

s.t. Ass. 2
$$0.3W_2 + 0.5P_2 \le 10$$

Mol. 2 $W_2 + P_2 \le 21$
Ste. 2 $-S + 1.5W_2 + P_2 \le 0$
W.d. 2 $W_2 \le 15$
P.d. 2 $P_2 \le 16$
Obj. 2 $-Z_2 + 160W_2 + 100P_2 = 0$

Formulation of the two-stage problem

and

s.t. Ass. 3
$$0.3W_3 + 0.5P_3 \le 8$$

Mol. 3 $W_3 + P_3 \le 21$
Ste. 3 $-S + 1.5W_3 + P_3 \le 0$
W.d. 3 $W_3 \le 15$
P.d. 3 $P_3 \le 16$
Obj. 3 $-Z_3 + 90W_3 + 100P_3 = 0$

Formulation of the two-stage problem

and finally

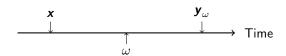
s.t. Ass.
$$4$$
 $0.3W_4 + 0.5P_4 \le 10$
Mol. 4 $W_4 + P_4 \le 21$
Ste. 4 $-S + 1.5W_4 + P_4 \le 0$
W.d. 4 $W_4 \le 15$
P.d. 4 $P_4 \le 16$
Obj. 4 $-Z_4 + 90W_4 + 100P_4 = 0$
 $S \ge 0$; $W_i, P_i \ge 0, 1 \le i \le 4$

Formulation of the two-stage problem

Solution: S = 27,250lb.

	$ W_i $	P_i
1	15,000	4,750
2	15,000	4,750
3	12,500	8,500
4	5,000	16,000

Two-stage optimization problems



• First stage decisions: x:

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq 0$$

- Random scenarios indexed by $\omega=1,\ldots,K.$ scenario ω has probability α_ω
- Second stage decisions: \mathbf{y}_{ω} : $\omega = 1, \dots, K$.

$$m{B}_{\omega}m{x} + m{D}_{\omega}m{y}_{\omega} = m{d}_{\omega}, \quad \ m{y}_{\omega} \geq 0$$



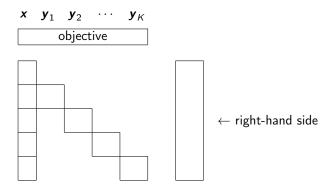
Two-stage optimization problems

- Objective: Find x, y_1, \dots, y_K so at to minimize the "expected cost":
- Formulation

• Note that even if the number of scenarios *K* is moderate, this formulation can have a lot of variables, but it has a nice structure ...

Two-stage optimization problems

Structure of the formulation



Two-stage optimization problems - how to solve?



Observation:

- Given a first-stage decision x, the second-stage problems can all be solved separately.
- Given x and a realization ω , the optimal second-stage cost is

$$egin{aligned} oldsymbol{z}_{\omega}(oldsymbol{x}) := \min & oldsymbol{f}^{T} oldsymbol{y}_{\omega} \ & ext{s.t.} & oldsymbol{D}_{\omega} oldsymbol{y}_{\omega} = oldsymbol{d}_{\omega} - oldsymbol{B}_{\omega} oldsymbol{x} \ & oldsymbol{y}_{\omega} \geq 0 \end{aligned}$$

 \bullet To find the optimal first-stage decision x we need to solve

min
$$c^T x + \sum_{\omega=1}^K \alpha_{\omega} z_{\omega}(x)$$

s.t. $Ax = b$
 $x \ge 0$



Two-stage optimization problems - Idea

Use the dual formulation of the second-stage problems:

Strong duality

$$z_{\omega}(\mathbf{x}) = \max_{\mathbf{p}_{\omega}^{T}} (\mathbf{d}_{\omega} - \mathbf{B}_{\omega} \mathbf{x})$$

s.t. $\mathbf{p}_{\omega}^{T} \mathbf{D}_{\omega} \leq \mathbf{f}^{T}$

holds when dual feasible set $\mathcal{P}_{\omega} := \left\{ \boldsymbol{p} : \boldsymbol{p}^T \boldsymbol{D}_{\omega} \leq \boldsymbol{f}^T \right\}$ is nonempty and bounded.

• In that case, we also know that

$$z_{\omega}(\mathbf{x}) = \max_{i} \; \mathbf{p}_{\omega,i}^{T}(\mathbf{d}_{\omega} - \mathbf{B}_{\omega}\mathbf{x})$$

where ${m p}_{\omega,i}$ are the corners of the dual polytope $P_{\omega}.$



Two-stage optimization problems - Reformulation

Reformulation of the first-stage problem:

min
$$c^T x + \sum_{\omega=1}^K \alpha_{\omega} z_{\omega}$$

s.t. $Ax = b$
 $p_{\omega,i}^T (d_{\omega} - B_{\omega} x) \le z_{\omega} \quad \forall i, \omega$
 $x \ge 0$

Observation:

- In this reformulation the number of variables has been reduced significantly.
- Potential huge number of constraints ⇒ cutting plane methods.

4 D > 4 D > 4 E > 4 E > E 9 Q P

Two-stage optimization problems - Feasibility check and row generation

We can check feasibility (\bar{x}, \bar{z}) very easily by computing the second stage costs

$$egin{array}{lll} oldsymbol{z}_{\omega}(ar{oldsymbol{x}}) := \min & oldsymbol{f}^T oldsymbol{y}_{\omega} &= oldsymbol{max} & oldsymbol{p}_{\omega}^T oldsymbol{d}_{\omega} - oldsymbol{B}_{\omega} ar{oldsymbol{x}} & ext{s.t.} & oldsymbol{p}_{\omega}^T oldsymbol{c}_{\omega} - oldsymbol{B}_{\omega} ar{oldsymbol{x}} \end{pmatrix} \ & ext{s.t.} & oldsymbol{p}_{\omega}^T oldsymbol{D}_{\omega} \leq oldsymbol{f}^T \ & oldsymbol{p}_{\omega}^T oldsymbol{D}_{\omega} \leq oldsymbol{f}^T \ & ext{s.t.} & oldsymbol{f}_{\omega} = oldsymbol{f}_{\omega} oldsymbol{f}_{\omega} = oldsymbol{f}_{\omega} oldsymbol{f}_{\omega} + oldsymbol{f}_{\omega} oldsymbol{f}_{\omega} = oldsymbol{f}_{\omega} + oldsymbol{f}_{\omega} oldsymbol{f}_{\omega} + oldsymbo$$

to get $z_{\omega}(\bar{x})$ and corresponding optimal dual p_{ω}^{\star} for all ω .

Possibilities:

- $\bar{z}_{\omega} \geq z_{\omega}(\bar{x})$ for all ω ; (\bar{x}, \bar{z}) is feasible.
- $\bar{z}_{\omega} < z_{\omega}(\bar{x})$ for some ω ; We add a violating constraint by demanding

$$z_{\omega} \geq (\boldsymbol{p}_{\omega}^{\star})^{T}(\boldsymbol{d}_{\omega} - \boldsymbol{B}_{\omega}\bar{\boldsymbol{x}}) \ \ell$$



Two-stage optimization problems - A complete iteration

Solve relaxed master problem

$$\begin{aligned} & \min \quad \boldsymbol{c}^{T}\boldsymbol{x} + \sum_{\omega=1}^{K} \alpha_{\omega} z_{\omega} \\ & \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \quad \boldsymbol{p}_{\omega,i}^{T}(\boldsymbol{d}_{\omega} - \boldsymbol{B}_{\omega}\boldsymbol{x}) \leq z_{\omega} \quad \forall (i,\omega) \in \mathcal{C}_{\ell} \\ & \quad \boldsymbol{x} \geq 0 \end{aligned}$$

The sets C_{ℓ} contain some but not all constraints. Primal solution $(\mathbf{x}_{\ell}, \mathbf{z}_{\ell})$.

32 / 33

Large Scale 6.255J @ 2021 (MIT)

Two-stage optimization problems - A complete iteration

② Check feasibility $(\mathbf{x}_{\ell}, \mathbf{z}_{\ell})$ by checking for all ω that

$$z_{\ell,\omega} \ge z_{\omega}(\mathbf{x}_{\ell}) := \min \quad \mathbf{f}^{\mathsf{T}} \mathbf{y}_{\omega}$$

s.t. $\mathbf{D}_{\omega} \mathbf{y}_{\omega} = \mathbf{d}_{\omega} - \mathbf{B}_{\omega} \mathbf{x}_{\ell}$
 $\mathbf{y}_{\omega} \ge 0$

with dual optimal basic solution $\boldsymbol{p}_{\ell,\omega}$.

- 3 If so then solution (x_{ℓ}, z_{ℓ}) is optimal
- **4** Otherwise select a violating ω_{ℓ} .
- Optimal dual solution $\boldsymbol{p}_{\omega_{\ell}}^{\star} = \boldsymbol{p}_{\omega_{\ell}, i_{\ell}}$ is a corner of the dual feasible set. \Rightarrow Add constraint : $\mathcal{C}_{\ell+1} = \mathcal{C}_{\ell} \cup \{(\omega_{\ell}, i_{\ell})\}$.
- Resolve updated relaxed master problem

