M2 ORO: Advanced Integer Programming
Solution Final Exam – 1st session

november 15, 2010

duration: 1h30.

documents: lecture notes are authorized. No book, no book copy.

grades: 3 problems of respectively 5+10+5 points each = 20 points.

Notations:

$\mathbb{Z}, \mathbb{Z}_+, \mathbb{Z}_{+*}$	the sets of integer, non-negative integer, and positive integer numbers
$\mathbb{R}, \mathbb{R}_+, \mathbb{R}_{+*}$	the sets of real, non-negative real, and positive real numbers

---

1 Modeling and decomposition

Problem 1 The Graph Clustering Problem.
Consider a complete graph $G = (V, E)$, *a constant* $K \in \mathbb{Z}_{+*}$, *a cost* $c_e > 0$ *for each edge* $e \in E$, *a weight* $d_i \geqslant 0$ *for each node* $i \in V$, *and a cluster capacity* C *with* $\min_{i \in V} d_i \leqslant C < \sum_{i \in V} d_i$.

A ***capacitated cluster*** *of* G *is a (possibly empty) subset of nodes satisfying the property that the sum of the node weights does not exceed the capacity* C*; an edge is included in a cluster if it joins two nodes within the cluster.*

The problem is to split the node set V *into* K *capacitated clusters such that any node belongs to at most one cluster and the sum of the costs of the edges included in any clusters is maximized.*

Question 1 ***(5 points).***

Q1.1. Model this problem as a Binary Linear Program.

Q1.2. Identify, if they exist, the redundant constraints and variables in this formulation.

Q1.3. Reformulate the problem as a set-packing problem and derive a second Binary Linear Program formulation for the graph clustering problem.

Q1.4. The LP-relaxation of this second program can be solved using a column-generation approach. Formulate the pricing problem involved in such a decomposition as an Integer Linear Program.

1.

$$
\begin{aligned}
(GC): z \ = \ \max & \sum_{k=1}^{K} \sum_{e \in E} c_e y_e^k \\
\text{s.t.} & \sum_{k=1}^{K} x_i^k \leqslant 1 && \forall i \in V, \\
& \sum_{i \in V} d_i x_i^k \leqslant C && \forall k = 1, \dots, K, \\
& y_e^k \leqslant x_i^k && \forall e = \in E, i \in e, k = 1, \dots, K, \\
& y_e^k \geqslant x_i^k + x_j^k - 1 && \forall e = (i,j) \in E, k = 1, \dots, K, \\
& x_i^k \in \{0,1\} && \forall i \in V, k = 1, \dots, K, \\
& y_e^k \in \{0,1\} && \forall e \in E, k = 1, \dots, K.
\end{aligned}
$$

where $y_e^k = 1$ if e is included in the k-th cluster and $x_i^k = 1$ if i belongs to the k-th cluster.

2. consider any optimal solution (x^*, y^*) of the BIP (GC') obtained from (GC) by dropping the set of constraints $y_e^k \geqslant x_i^k + x_j^k - 1$, $\forall (e,k)$. Such a solution exists since (GC') is feasible (consider solution $x = 0, y = 0$) and bounded (by $\sum_{e \in E} c_e$). Assume that this solution does not satisfy one of the relaxed constraint, i.e. there exists $e = (i,j) \in E$ and $k \in [1, K]$ such that $y_e^{k*} = 0$ and $x_i^{k*} = x_j^{k*} = 1$. Then this variable y_e^{k*} can be set to 1, without violating the feasibilty of the solution (x^*, y^*) in (GC'), but strictly improving its cost by $c_e > 0$. It is absurd since the initial solution was optimal. Hence, any optimal solution of (GC') is feasible in (GC). It is even optimal in (GC) since (GC') is a relaxation of (GC). As a consequence, constraints $y_e^k \geqslant x_i^k + x_j^k - 1$ are redundant.

3. Let S be the set of all non-empty capacitated clusters of G: $S = \{s \subset V \mid s \neq \emptyset, \sum_{i \in s} d_i \leqslant C\}$ and consider $c_s = \sum c_e \mid e \subseteq s\}$ the cost of element $s \in S$. Then the problem is to find at most K elements of S, such that each node $i \in V$ belongs to at most one selected element of S and the sum of the element costs is maximized. For all $i \in V$ and $s \in S$, let $\delta_{is} = 1$ if $i \in s$, and 0 otherwise. The problem can be formulated using binary variables $x_s = 1$ iff $s \in S$ is selected in the solution:

$$\begin{aligned} \max \ & \sum_{s \in S} c_s x_s \\ \text{s.t.} \ & \sum_{s \in S} x_s \leqslant K, \\ & \sum_{s \in S} \delta_{is} x_s \leqslant 1 && \forall i \in V, \\ & x_s \in \{0,1\} && \forall s \in S. \end{aligned}$$

4. the pricing problem is to find a violated dual constraint, given an optimal solution of the master program. Let $\mu \geqslant 0$ and $\lambda_i \geqslant 0 \ \forall i \in V$ form a dual optimum solution of the master then we look for a set $s \in S$ such that $\mu + \sum_{i \in V} \delta_{is} \lambda_i < c_s$. The pricing problem is then to solve the following knapsack problem:

$$\begin{aligned} \max \ & \mu + \sum_{i \in V} z_i \lambda_i \\ \text{s.t.} \ & \sum_{i \in V} d_i z_i \leqslant C, \\ & z_i \in \{0,1\} && \forall i \in V. \end{aligned}$$

If its optimum value is lower than c_s then its solution is a possible entering column, otherwise the column generation process stops as the solution of the master is optimum for the LP-relaxation.

---

2 Cutting-planes

Problem 2 Uncapacited Lot-Sizing Problem (ULS).
The ULS problem is to decide a production plan for a n-period horizon for a single product, given:

- $f_t \in \mathbb{Z}_{+*}$ *the fixed cost of producing in period* t
- $p_t \in \mathbb{Z}_{+*}$ *the unit production cost in period* t
- $h_t \in \mathbb{Z}_{+*}$ *the unit storage cost in period* t
- $d_t \in \mathbb{Z}_{+*}$ *the demand in period* t.

The production plan must satisfy the demand with a minimum cost.
We know two formulations for the ULS problem as Mixed Integer Linear Programs, the aggregated model (P)*:*

$$\begin{aligned} (P): \ \min \ & \sum_{t=1}^{n} f_t y_t + \sum_{t=1}^{n} p_t x_t + \sum_{t=1}^{n} h_t s_t \\ \textit{s.t.} \ & s_{t-1} + x_t = d_t + s_t && t = 1, \dots, n \\ & x_t \leqslant M_t y_t && t = 1, \dots, n \\ & y_t \in \{0,1\} && t = 1, \dots, n \\ & s_t, x_t \geqslant 0 && t = 1, \dots, n \\ & s_0 = 0 \end{aligned}$$

and the extended model (E)*:*

$$\begin{aligned} (E): \ \min \ & \sum_{t=1}^{n} f_t y_t + \sum_{i=1}^{n} \sum_{t=i}^{n} p_i z_{it} + \sum_{i=1}^{n} \sum_{t=i+1}^{n} \sum_{j=i}^{t-1} h_j z_{it} \\ \textit{s.t.} \ & \sum_{i=1}^{t} z_{it} = d_t && t = 1, \dots, n \\ & z_{it} \leqslant d_t y_i && i = 1, \dots, n; t = i, \dots, n \\ & y_t \in \{0,1\} && t = 1, \dots, n \\ & z_{it} \geqslant 0 && i = 1, \dots, n; t = i, \dots, n \end{aligned}$$

We know that the second formulation (E) *is ideal but not the first one* (P)*. However* (E) *contains a quadratic number of variables and constraints. When the number of periods* n *increases, the size of this model may be too large to solve it quickly. We might therefore look for a cutting-plane approach based on the first formulation* (P)*, which only contains a linear number of variables and constraints.*

Question 2 ***(10 points).***

Q2.1. Is ULS an easy problem ? Explain how and why.

Q2.2. Find the tightest possible values M_t in model (P).

Q2.3. Reformulate the LP-relaxation of (P) as a Linear Program with only $2n$ non-negative variables and n constraints (*hint: replace the* y *variables*). Can we also remove the x_t variables, in this formulation, replacing them in the objective by the expression $d_t + s_t - s_{t-1}$, hence obtaining a linear program on the non-negative s variables alone ?

Q2.4. The demands d_t at any period t are positive ($d_t > 0$) and must be satisfied. Infer that one variable y_t of (P) can be fixed for some period t.

Q2.5. Infer, for any period t, a lower bound of the entering stock s_{t-1} whenever no production occurs at period t. Derive from this, a valid inequality C_t for (P) linking the s_{t-1} and y_t variables.

Q2.6. Show that the following inequality is valid for (P) for any period t:

$$D_t: \quad x_t \leqslant d_t y_t + s_t.$$

Compare the strength of the two constraints C_t and D_t.

Q2.7.(*) Show that the following constraints are all valid inequalities for (P):

$$F_{(l,S)}: \quad \sum_{i \in S} x_i \leqslant \sum_{i \in S} (\sum_{t=i}^{l} d_t) y_i + s_l, \quad \forall l \in \{1, \dots, n\}, \forall S \subseteq \{1, \dots, l\}.$$

(*proof by induction on the size of* S.)

Q2.8. Propose a separation algorithm for the class of valid inequalities $F_{(l,S)}$, returning, for any fixed period l, the cut $F_{(l,S)}$ the most violated by a given optimal fractional solution of (P) (if such violated inequalities exist). What is its worst-case time complexity ?

Epilog: *the following Linear Program is an ideal formulation of* ULS*:*

$$
\begin{aligned}
(L):\ \min\ & \sum_{t=1}^{n} f_t y_t + \sum_{t=1}^{n} p_t x_t + \sum_{t=1}^{n} h_t s_t \\
\text{s.t. } & s_{t-1} + x_t = d_t + s_t && t = 1,\dots,n \\
& x_t \leqslant (\sum_{k=t}^{n} d_k) y_t && t = 1,\dots,n \\
& \sum_{i\in S} x_i \leqslant \sum_{i\in S} (\sum_{t=i}^{l} d_t) y_i + s_l && \forall l = 1,\dots,n, \forall S \subseteq \{1,\dots,l\} \\
& y_t \leqslant 1 && t = 1,\dots,n \\
& s_t, x_t, y_t \geqslant 0 && t = 1,\dots,n \\
& s_0 = 0, y_1 = 1
\end{aligned}
$$

1. (E) is an ideal formulation of ULS of polynomial size. As a consequence, the LP-relaxation of (E) can be solved in polynomial time and leads to an integer optimal solution of ULS.
2. $M_t = \sum_{i=t}^{n} d_i$ the remaining demand to satisfy.
3. In the LP-relaxation of (P), the y variables can now take fractionnal values between 0 and 1. These variables are only constrained, each one individually by $y_t \geqslant x_t/M_t$, and must be minimized. Then in any optimal fractionnal solution, $y_t = x_t/M_t$. Note that constraint $x_t \leqslant M_t$ (or $y_t \leqslant 1$) is redundant with the minimization objective. Hence, the LP-relaxation of (P) can be reformulated as:

$$
\begin{aligned}
(\bar{P}'):\ \min\ & \sum_{t=1}^{n} (\frac{f_t}{M_t} + p_t) x_t + \sum_{t=1}^{n} h_t s_t \\
\text{s.t. } & s_{t-1} + x_t = d_t + s_t && t = 1,\dots,n \\
& s_t, x_t \geqslant 0 && t = 1,\dots,n \\
& s_0 = 0
\end{aligned}
$$

   The x variables can be substituted using expression $x_t = d_t + s_t - s_{t-1}$ but one should not miss constraint $x_t \geqslant 0$ that becomes $s_{t-1} \leqslant d_t + s_t$. Conversely, the s variables can be substitued in the objective and in constraint $s_t \geqslant 0$ by $s_t = \sum_{i=1}^{t} (x_i - d_i)$.
4. The demand at time 1 must be satisfied and $s_0 = 0$: $x_1 = d_1 + s_1 \geqslant d_1 > 0$, then $y_1 = 1$.
5. if $y_t = 0$ then $x_t = 0$ and $s_{t-1} = d_t + s_t \geqslant d_t$. This logical constraint can be modeled as $s_{t-1} \geqslant d_t(1 - y_t)$.
6. if $y_t = 0$ then $x_t = 0 \leqslant s_t$ and if $y_t = 1$ then $x_t = d_t + s_t - s_{t-1} \leqslant d_t + s_t$. This constraint is equivalent to the previous one: $x_t = d_t + s_t - s_{t-1} \geqslant d_t + s_t - d_t(1 - y_t) = d_t y_t + s_t$.
7. if $|S| = 0$ the constraint is trivial $0 \leqslant s_l$ for all l. Let the constraint be valid for all l and $|S| \leqslant k$, with $k \geqslant 0$ and consider some period $l = 1,\dots,n$, and some subset $S \subseteq \{1,\dots,l\}$ of size $k+1$. Then, let j be the maximum element of S and $S' = S \setminus \{j\}$ then $\sum_{i\in S} x_i = x_j + \sum_{i\in S'} x_i \leqslant x_j + \sum_{i\in S'} (\sum_{t=i}^{l} d_t) y_i + s_l$. If $x_j = 0$ then
8. The algorithm is to find, given a fractionnal solution $(\bar{x}, \bar{s})$ and a time period l a subset $S \subseteq \{1,\dots,n\}$ such that $w_S = \sum_{i\in S} \bar{x}_i (1 - (\sum_{t=i}^{l} d_t)/M_i) - \bar{s}_l$ is positive and maximal. This is equivalent to solve a 0-1 Knapsack problem with l items with weights $w_i = \bar{x}_i (1 - (\sum_{t=i}^{l} d_t)/M_i)$

Problem 3 Capacitated Lot-Sizing Problem (CLS).
CLS *is a variant of* ULS *where the number of items produced at any period* t *must not exceed a given capacity* $c_t \in \mathbb{Z}_{+*}$.

Question 3 ***(5 points).***

Q3.1. Model CLS as a Mixed Integer Linear Program based on the aggregate formulation (E) of ULS.

Q3.2. Model CLS as a Mixed Integer Linear Program based on the aggregate formulation (P) of ULS.

Q3.3. Which constraints of (L) are still valid for this formulation of CLS ?