IE 631 Integer Programming **KAIST, Spring 2023**
Lecture #20: Uncapacitated facility location annd the subgradient algorithm May 11, 2023
Lecturer: Dabeen Lee

1 Outline

In this lecture, we study

- the uncapacitated facility location problem.
- subgradient algorithm for the Lagrangian dual.

2 Uncapacitated facility location

There are m customers, and each customer $i \in [m]$ has demand d_i. There are d locations for building facilities. Each location $j \in [d]$ requires fixed annual operating cost f_j. Let c_{ij} is the cost of transporting one unit of item from location j to customer i. The (uncapacitated) facility location problem is to determine where to build facilities so as to minimize the total cost while satisfying the customer demands.

To formulate the problem, we introduce variable x_j to indicate whether we build a facility in location j or not.

$$x_j = \begin{cases} 1, & \text{if we build a facility in } j, \\ 0, & \text{otherwise.} \end{cases}$$

Let y_{ij} be the fraction of demand d_i of customer i that is dealt with by facility in j. Then the problem can be formulated as

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} \sum_{j=1}^{d} c_{ij} d_i y_{ij} + \sum_{j=1}^{n} f_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} y_{ij} = 1, \quad i = 1, \ldots, m \\ & 0 \leq y_{ij} \leq x_j, \quad i = 1, \ldots, m, \ j = 1, \ldots, d \\ & x \in \{0,1\}^d. \end{aligned}$$

Theorem 20.1. *For every customer $i \in [m]$, it is optimal to satisfy its demand by a single fixed facility.*

Therefore, we may assume that y is binary, and the problem can be reformulated as

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} \sum_{j=1}^{d} c_{ij} d_i y_{ij} + \sum_{j=1}^{n} f_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} y_{ij} = 1, \quad i = 1, \ldots, m \\ & y_{ij} \leq x_j, \quad i = 1, \ldots, m, \ j = 1, \ldots, d \\ & y \in \{0,1\}^{m \times d}, \ x \in \{0,1\}^d. \end{aligned}$$

Here, the constraints

$$\sum_{j=1}^{n} y_{ij} = 1, \quad i = 1, \ldots, m$$

are complicating constraints. Given a multiplier vector $\lambda \in \mathbb{R}^m$, the associated Lagrangian relaxation is given by

$$\begin{aligned} z_{\text{LR}}(\lambda) \quad = \quad \min \quad & \sum_{i=1}^{m} \sum_{j=1}^{d} c_{ij} d_i y_{ij} + \sum_{j=1}^{n} f_j x_j + \sum_{i=1}^{m} \lambda_i (1 - \sum_{j=1}^{d} y_{ij}) \\ \text{s.t.} \quad & y_{ij} \le x_j, \quad i = 1, \ldots, m, \ j = 1, \ldots, d \\ & y \in \{0,1\}^{m \times d}, \ x \in \{0,1\}^d. \end{aligned}$$

We may rewrite this as

$$\begin{aligned} z_{\text{LR}}(\lambda) \quad = \quad \min \quad & \sum_{i=1}^{m} \sum_{j=1}^{d} (c_{ij} d_i - \lambda_i) y_{ij} + \sum_{j=1}^{n} f_j x_j + \sum_{i=1}^{m} \lambda_i \\ \text{s.t.} \quad & y_{ij} \le x_j, \quad i = 1, \ldots, m, \ j = 1, \ldots, d \\ & y \in \{0,1\}^{m \times d}, \ x \in \{0,1\}^d. \end{aligned}$$

Then the corresponding Lagrangian dual is defined as

$$z_{\text{LD}} = \max\{z_{\text{LR}}(\lambda) : \ \lambda \in \mathbb{R}^m\}.$$

The advantage of working over the Lagrangian relaxation is that the constraints

$$y_{ij} \le x_j, \quad i = 1, \ldots, m, \ j = 1, \ldots, d$$

give rise to a totally unimodular constraint matrix. Therefore, it follows that

$$\begin{aligned} z_{\text{LR}}(\lambda) \quad = \quad \min \quad & \sum_{i=1}^{m} \sum_{j=1}^{d} (c_{ij} d_i - \lambda_i) y_{ij} + \sum_{j=1}^{n} f_j x_j + \sum_{i=1}^{m} \lambda_i \\ \text{s.t.} \quad & y_{ij} \le x_j, \quad i = 1, \ldots, m, \ j = 1, \ldots, d \\ & 0 \le y_{ij}, x_i \le 1, \quad i = 1, \ldots, m, \ j = 1, \ldots, d. \end{aligned}$$

3 Subgradient algorithm for the Lagrangian dual

Given a convex function $f : \mathbb{R}^m \to \mathbb{R}$ and a fixed point $x \in \text{dom}(f)$, the **subdifferential** of f at x is defined as

$$\partial f(x) = \left\{ g : \ f(y) \ge f(x) + g^\top (y - x) \ \ \forall y \in \text{dom}(f) \right\}.$$

Here, any $g \in \partial f(x)$ is called a **subgradient** of f at x.

Conversely, the subdifferential is the set of subgradients. If function f is differentiable at x, then we have $\partial f(x) = \{\nabla f(x)\}$, and therefore, the subdifferential reduces to the gradient. In contrast, a non-differentiable function may have more than one subgradient. Moreover, note that for any subgradient g at x, $f(x) + g^\top (y - x)$ provies a lower approximation of the function f.

Recall that for a differentiable univariate function f, the gradient of f at some point x is the slope of the line tangent to f at x. We have a similar geometric intuition for subgradients. Consider the the absolute value function $f(x) = |x|$ over $x \in \mathbb{R}$, which is not differentiable at $x = 0$.

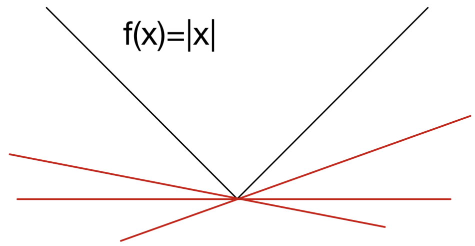

Figure 20.1: Subgradients of $f(x) = |x|$ at $x = 0$

Proposition 20.2. *Let $\bar{x}$ be an optimal solution to the Lagrangian relaxation with respect to $\bar{\lambda}$. Then $f - E\bar{x}$ is a subgradient of z_{LR} at $\bar{\lambda}$.*

Proof. For any $\lambda \geq 0$, we have

$$\begin{aligned} z_{\mathrm{LR}}(\lambda) &\geq c^\top \bar{x} + \lambda^\top (f - E\bar{x}) \\ &= c^\top \bar{x} + \bar{\lambda}^\top (f - E\bar{x}) + (\lambda - \bar{\lambda})^\top (f - E\bar{x}) \\ &= z_{\mathrm{LR}}(\bar{\lambda}) + (\lambda - \bar{\lambda})^\top (f - E\bar{x}). \end{aligned}$$

Therefore, $f - A\bar{x}$ is a subgradient of z_{LR} at $\bar{\lambda}$, as required. □

Recall that solving the Lagrangian dual problem is equivalent to minimizing the Lagrangian relaxation function. We may apply the **subgradient method**.

Algorithm 1 Subgradient method

Initialize $\lambda_1 \in \mathrm{dom}(z_{\mathrm{LR}})$.
for $t = 1, \ldots, T$ **do**
  Solve the Lagrangian relaxation with respect to λ_t.
  Obtain $f - Ex_t$ where x_t is an optimal solution to the Lagrangian relaxation.
  Update $\lambda_{t+1} = [\lambda_t - \eta_t(f - Ex_t)]_+$ for some step size $\eta_t > 0$.
end for

In Algorithm 1, we have

$$\lambda_{t+1} = [\lambda_t - \eta_t(f - Ex_t)]_+ = \mathrm{proj}_{\mathbb{R}^m_+} (\lambda_t - \eta_t(f - Ex_t)).$$

Basically, $[\lambda_t - \eta_t(f - Ex_t)]_+$ is obtained from $\lambda_t - \eta_t(f - Ex_t)$ after replacing each of its negative components by 0.

How do we set up the step sizes $\{\eta_t\}_{t=1}^T$?

Theorem 20.3 (Poljak, 1967)**.** *Suppose that*

$$\sum_{t=1}^{\infty} \eta_t = +\infty \quad \textit{and} \quad \lim_{t \to \infty} \eta_t = 0.$$

Then the subgradient method (Algorithm 1) converges to z_{LD}, the Lagrangian dual value.

An example is $\eta_t = 1/\sqrt{t}$. Then the convergence rate is $O(1/\sqrt{T})$, which is the optimal rate for general convex functions. However, for the Lagrangian dual of a mixed-integer program, it is often the case that the subgradient method with a Poljak rule converges very slow.

An alternate choice is a geometric series, i.e.,

$$\eta_t = \eta_0 \rho^t, \quad \rho \in (0, 1).$$

With a geometric series, the subgradient method converges very fast, but not necessarily to the optimum. **However, this may not be a critical issue in the context of branch-and-bound.**