Advanced Algorithms

南京大学

尹一通
LP-based Algorithms

• LP rounding:
  • Relax the integer program to LP;
  • round the optimal LP solution to a nearby feasible integral solution.

• The primal-dual schema:
  • Find a pair of solutions to the primal and dual programs which are close to each other.
LP Duality

**Primal:**
\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad Ax \geq b \\
& \quad x \geq 0
\end{align*}
\]

**Dual:**
\[
\begin{align*}
\max & \quad b^T y \\
\text{s.t.} & \quad y^T A \leq c^T \\
& \quad y \geq 0
\end{align*}
\]

**Monogamy:** dual(dual(LP)) = LP

**Weak Duality:**
\[
\forall \text{ feasible primal solution } x \text{ and dual solution } y \\
y^T b \leq y^T A x \leq c^T x
\]
LP Duality

Primal:
\[
\begin{align*}
\text{min} & \quad c^T x \\ 
\text{s.t.} & \quad Ax \geq b \\ 
& \quad x \geq 0
\end{align*}
\]

Dual:
\[
\begin{align*}
\text{max} & \quad b^T y \\ 
\text{s.t.} & \quad y^T A \leq c^T \\ 
& \quad y \geq 0
\end{align*}
\]

Weak Duality Theorem:
\[
\forall \text{ feasible primal solution } x \text{ and dual solution } y \quad y^T b \leq c^T x
\]
LP Duality

Primal:
\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad Ax \geq b \\
& \quad x \geq 0
\end{align*}
\]

Dual:
\[
\begin{align*}
\max & \quad b^T y \\
\text{s.t.} & \quad y^T A \leq c^T \\
& \quad y \geq 0
\end{align*}
\]

Strong Duality Theorem:
Primal LP has finite optimal solution \( x^* \)
\[\text{iff}\] dual LP has finite optimal solution \( y^* \).
\[
y^{*T}b = c^Tx^*
\]
Max-Flow

digraph: \( D(V,E) \)  
source: \( s \)  
sink: \( t \)

capacity: \( c : E \rightarrow \mathbb{R}^+ \)

\[
\begin{align*}
\text{max} & \quad \sum_{u:(s,u)\in E} f_{su} \\
\text{s.t.} & \quad 0 \leq f_{uv} \leq c_{uv} \\
& \quad \forall (u, v) \in E \\
\text{for} \quad & \quad \sum_{w:(w,u)\in E} f_{wu} - \sum_{v:(u,v)\in E} f_{uv} = 0 \\
& \quad \forall u \in V \setminus \{s, t\}
\end{align*}
\]
Max-Flow

digraph: $D(V,E)$  source: $s$  sink: $t$
capacity: $c : E \rightarrow \mathbb{R}^+$

$max$  $f_{ts}$

$d_{uv}$  s.t.  $0 \leq f_{uv} \leq c_{uv}$

$p_u \quad \sum_{w:(w,u) \in E} f_{wu} - \sum_{v:(u,v) \in E} f_{uv} \leq 0 \quad \forall u \in V$

$\forall (u, v) \in E$
Dual-LP

digraph: \( D(V,E) \)
source: \( s \)
sink: \( t \)
capacity: \( c : E \rightarrow \mathbb{R}^+ \)

\[
\min \sum_{(u,v) \in E} c_{uv} d_{uv}
\]

s.t. \( d_{uv} - p_u + p_v \geq 0 \)
\( p_s - p_t \geq 1 \)
\( d_{uv} \geq 0, \ p_u \geq 0 \)
\( \forall (u, v) \in E \)
\( \forall (u, v) \in E \)
\( \forall u \in V \)
Dual-LP

digraph: $D(V,E)$

capacity: $c: E \rightarrow \mathbb{R}^+$

min $\sum_{(u,v) \in E} c_{uv}d_{uv}$

s.t. $d_{uv} - p_u + p_v \geq 0$

$s$ $t$

$\forall (u, v) \in E$

$p_s - p_t \geq 1$

$d_{uv}, p_u \in \{0, 1\}$

$\forall (u, v) \in E$ $\forall u \in V$
Integral Polytope

Integral polyhedron:
all vertices are integral

OPT for IP =
OPT for LP-relaxation

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad Ax \geq b \\
& \quad x \in \mathbb{Z}^n \\
\end{align*}
\]

max-flow-min-cut ⇐ Strong Duality + Integrality of Cut Polytope
LP Duality

Primal:

\[
\begin{align*}
\text{min} & \quad c^T x \\
\text{s.t.} & \quad Ax \geq b \\
& \quad x \geq 0
\end{align*}
\]

Dual:

\[
\begin{align*}
\text{max} & \quad b^T y \\
\text{s.t.} & \quad y^T A \leq c^T \\
& \quad y \geq 0
\end{align*}
\]

Strong Duality Theorem:

Primal LP has finite optimal solution \( x^* \)
iff dual LP has finite optimal solution \( y^* \), and

\[
y^{*T} b = c^T x^*
\]
Primal: \( \min \ c^T x \) \quad \text{Dual: } \max \ b^T y \\
\text{s.t. } A x \geq b \quad \text{s.t. } y^T A \leq c^T \\
x \geq 0 \quad y \geq 0

\forall \text{ feasible primal solution } x \text{ and dual solution } y \\
\quad y^T b \leq y^T A x \leq c^T x

\textbf{Strong Duality Theorem} \quad x \text{ and } y \text{ are both optimal iff} \\
\quad y^T b = y^T A x = c^T x

\forall i: \text{ either } A_i \cdot x = b_i \text{ or } y_i = 0 \\
\forall j: \text{ either } y^T A \cdot j = c_j \text{ or } x_j = 0
Complementary Slackness

Primal: \[ \min c^T x \quad \text{s.t.} \quad Ax \geq b \quad x \geq 0 \]

Dual: \[ \max b^T y \quad \text{s.t.} \quad y^T A \leq c^T \quad y \geq 0 \]

Complementary Slackness Conditions:

\forall \text{ feasible primal solution } x \text{ and dual solution } y \text{, } x \text{ and } y \text{ are both optimal iff:}

\forall i: \text{ either } A_i \cdot x = b_i \text{ or } y_i = 0

\forall j: \text{ either } y^T A \cdot j = c_j \text{ or } x_j = 0
**Complementary Slackness**

Primal: \[
\begin{align*}
\text{min} & \quad c^T x \\
\text{s.t.} & \quad Ax \geq b \\
& \quad x \geq 0
\end{align*}
\]

Dual: \[
\begin{align*}
\text{max} & \quad b^T y \\
\text{s.t.} & \quad y^T A \leq c^T \\
& \quad y \geq 0
\end{align*}
\]

∀ feasible primal solution \( x \) and dual solution \( y \)

for \( \alpha, \beta \geq 1 \):

\( \forall i \): either \( A_i \cdot x \leq \alpha b_i \) or \( y_i = 0 \)

\( \forall j \): either \( y^T A \cdot j \geq c_j / \beta \) or \( x_j = 0 \)

\[c^T x \leq \alpha \beta \ b^T y \leq \alpha \beta \ \text{OPT}_{LP}\]

\[
\sum_{j=1}^{n} c_j x_j \leq \sum_{j=1}^{n} \left( \beta \sum_{i=1}^{m} a_{ij} y_i \right) x_j = \beta \sum_{i=1}^{m} \left( \sum_{j=1}^{n} a_{ij} x_j \right) y_i \leq \alpha \beta \sum_{i=1}^{m} b_i y_i
\]
Primal-Dual Schema

Primal IP: \[ \min \ c^T x \]
\[ \text{s.t.} \quad Ax \geq b \]
\[ x \in \mathbb{Z}_{\geq 0} \]

Dual LP-relax: \[ \max \ b^T y \]
\[ \text{s.t.} \quad y^T A \leq c^T \]
\[ y \geq 0 \]

Find a primal integral solution \( x \) and a dual solution \( y \)

for \( \alpha, \beta \geq 1 \):

\( \forall i \): either \( A_i \cdot x \leq \alpha b_i \) or \( y_i = 0 \)
\( \forall j \): either \( y^T A \cdot j \geq c_j / \beta \) or \( x_j = 0 \)

\[ c^T x \leq \alpha \beta \ b^T y \leq \alpha \beta \ \text{OPT}_{LP} \leq \alpha \beta \ \text{OPT}_{IP} \]
The Primal-Dual Schema

• Write down an LP-relaxation and its dual.

\[
\begin{align*}
\min & \quad c^T x \\
\text{s.t.} & \quad A x \geq b \\
& \quad x \in \mathbb{Z}_{\geq 0}
\end{align*}
\quad \text{and}
\quad \begin{align*}
\max & \quad b^T y \\
\text{s.t.} & \quad y^T A \leq c^T \\
& \quad y \geq 0
\end{align*}
\]

• Start with a **primal infeasible** solution \( x \) and a **dual feasible** solution \( y \) (usually \( x=0, y=0 \)).

• Raise \( x \) and \( y \) until \( x \) is feasible:
  • raise \( y \) until some dual constraints gets **tight** \( y^T A \cdot j = c_j \);
  • raise \( x_j \) (integrally) corresponding to the **tight** dual constraints.

• Show the **complementary slackness conditions**:

\[
\begin{align*}
\forall i: \text{ either } A_i \cdot x & \leq \alpha b_i \text{ or } y_i = 0 \\
\forall j: \text{ either } y^T A \cdot j & = c_j \text{ or } x_j = 0
\end{align*}
\]

\[c^T x \leq \alpha b^T y \leq \alpha \text{ OPT}\]
Integrality Gap

LP relaxation of vertex cover: given $G(V,E)$,

- **minimize** $\sum_{v \in V} x_v$
- **subject to** $\sum_{v \in e} x_v \geq 1$, $e \in E$
- $x_v \in \{0, 1\}$, $v \in V$
- $x_v \in [0, 1]$

Integrality gap = $\sup_I \frac{\text{OPT}(I)}{\text{OPT}_{LP}(I)}$

For the LP relaxation of vertex cover: integrality gap = 2
Facility Location

hospitals in Nanjing
Facility Location

**Instance:** set $F$ of facilities; set $C$ of clients; facility opening costs $f: F \rightarrow [0, \infty)$; connection costs $c: F \times C \rightarrow [0, \infty)$;

Find a subset $I \subseteq F$ of **opening** facilities and a way $\phi: C \rightarrow I$ of **connecting** all clients to them such that the total cost $\sum_{j \in C} c_{\phi(j), j} + \sum_{i \in I} f_i$ is minimized.

- uncapsicitated facility location;
- **NP**-hard; **AP** (**Approximation Preserving**) - reduction from Set Cover;
- [Dinur, Steuer 2014] no poly-time $(1-o(1))\ln n$-approx. algorithm unless **NP** = **P**.
**Metric Facility Location**

**Instance:** set $F$ of facilities; set $C$ of clients; facility opening costs $f: F \rightarrow [0, \infty)$; connection metric $d: F \times C \rightarrow [0, \infty)$; Find a subset $I \subseteq F$ of opening facilities and a way $\phi: C \rightarrow I$ of connecting all clients to them such that the total cost $\sum_{j \in C} d_{\phi(j), j} + \sum_{i \in I} f_i$ is minimized.

**Triangle Inequality:** $\forall i_1, i_2 \in F, \forall j_1, j_2 \in C$ 

$$d_{i_1j_1} + d_{i_2j_1} + d_{i_2j_2} \geq d_{i_1j_2}$$
**Instance:** set $F$ of facilities; set $C$ of clients; facility opening costs $f: F \rightarrow [0, \infty)$; connection metric $d: F \times C \rightarrow [0, \infty)$;

Find $\phi: C \rightarrow I \subseteq F$ to minimize $\sum_{j \in C} d_{\phi(j), j} + \sum_{i \in I} f_i$

**LP-relaxation:**

$$\min \sum_{i \in F, j \in C} d_{ij} x_{ij} + \sum_{i \in F} f_i y_i$$

s.t. $y_i - x_{ij} \geq 0$, \quad $\forall i \in F, j \in C$

$$\sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C$$

$$x_{ij}, y_i \geq 0, \quad x_{ij}, y_i \in \{0, 1\}, \quad \forall i \in F, j \in C$$
Primal:

\[
\begin{align*}
\text{min} & \quad \sum_{i \in F, j \in C} d_{ij} x_{ij} + \sum_{i \in F} f_i y_i \\
\text{s.t.} & \quad y_i - x_{ij} \geq 0, \quad \forall i \in F, j \in C \\
& \quad \sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C \\
& \quad x_{ij}, y_i \in \{0, 1\}, \quad \forall i \in F, j \in C
\end{align*}
\]

\(\alpha_j\): amount of value paid by client \(j\) to all facilities

\(\beta_{ij} \geq \alpha_j - d_{ij}\): payment to facility \(i\) by client \(j\) (after deduction)

complimentary slackness conditions:

(if ideally held)

\(x_{ij} = 1 \Rightarrow \alpha_j - \beta_{ij} = d_{ij}\); \quad \alpha_j > 0 \Rightarrow \sum_{i \in F} x_{ij} = 1\);

\(y_i = 1 \Rightarrow \sum_{j \in C} \beta_{ij} = f_i\); \quad \beta_{ij} > 0 \Rightarrow y_i = x_{ij}\);

Dual-relax:

\[
\begin{align*}
\text{max} & \quad \sum_{j \in C} \alpha_j \\
\text{s.t.} & \quad \alpha_j - \beta_{ij} \leq d_{ij}, \quad \forall i \in F, j \in C \\
& \quad \sum_{j \in C} \beta_{ij} \leq f_i, \quad \forall i \in F \\
& \quad \alpha_j, \beta_{ij} \geq 0, \quad \forall i \in F, j \in C
\end{align*}
\]

primal-relax.
\[
\min \sum_{i \in F, j \in C} d_{ij}x_{ij} + \sum_{i \in F} f_iy_i \\
\text{s.t.} \quad y_i - x_{ij} \geq 0, \quad \forall i \in F, j \in C \\
\sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C \\
\quad x_{ij}, y_i \in \{0, 1\}, \quad \forall i \in F, j \in C
\]

\[
\max \sum_{j \in C} \alpha_j \\
\text{s.t.} \quad \alpha_j - \beta_{ij} \leq d_{ij}, \quad \forall i \in F, j \in C \\
\sum_{j \in C} \beta_{ij} \leq f_i, \quad \forall i \in F \\
\quad \alpha_j, \beta_{ij} \geq 0, \quad \forall i \in F, j \in C
\]

Initially \( \alpha = 0, \beta = 0 \), no facility is open, no client is served; raise \( \alpha_j \) for all client \( j \) \textit{simultaneously} at a \textit{uniform continuous} rate:

- upon \( \alpha_j = d_{ij} \) for a closed facility \( i \): edge \((i,j)\) is \textit{tight}; fix \( \beta_{ij} = \alpha_j - d_{ij} \) as \( \alpha_j \) being raised;
- upon \( \sum_{j \in C} \beta_{ij} = f_i \): \textit{tentatively open} facility \( i \); connect all \textit{unserved} clients \( j \) with \textit{tight} \((i,j)\) to facility \( i \) and stop raising \( \alpha_j \); all such clients \( j \) are \textit{served} by facility \( i \);
- upon \( \alpha_j = d_{ij} \) for a \textit{tentatively open} facility \( i \): connect client \( j \) to facility \( i \) and stop raising \( \alpha_j \); such client \( j \) is \textit{served} by facility \( i \);
Initially $\alpha = 0$, $\beta = 0$, no facility is open, no client is served; raise $\alpha_j$ for all client $j$ simultaneously at a uniform continuous rate:

- upon $\alpha_j = d_{ij}$ for a closed facility $i$: edge $(i,j)$ is tight; fix $\beta_{ij} = \alpha_j - d_{ij}$ as $\alpha_j$ being raised;
- upon $\sum_{j \in C} \beta_{ij} = f_i$: tentatively open facility $i$; connect all unserved clients $j$ with tight $(i,j)$ to facility $i$ and stop raising $\alpha_j$; all such clients $j$ are served by facility $i$;
- upon $\alpha_j = d_{ij}$ for a tentatively open facility $i$: connect client $j$ to facility $i$ and stop raising $\alpha_j$; such client $j$ is served by facility $i$;

- The events that occur at the same time are processed in arbitrary order.
- Fully paid facilities are tentatively open: $\sum_{j \in C} \beta_{ij} = f_i$
- Eventually every client is served by one tentatively opening facilities.
- Tight edges to opening facilities should be connected: $\alpha_j - \beta_{ij} = d_{ij}$

A client may have tight edges to more than one facilities! (opened more facilities than actually needed)
Initially $\alpha = 0$, $\beta = 0$, no facility is open, no client is served; raise $\alpha_j$ for all client $j$ simultaneously at a uniform continuous rate:

- upon $\alpha_j = d_{ij}$ for a closed facility $i$: edge $(i,j)$ is tight; fix $\beta_{ij} = \alpha_j - d_{ij}$ as $\alpha_j$ being raised;
- upon $\sum_{j \in C} \beta_{ij} = f_i$: tentatively open facility $i$; connect all unserved clients $j$ with tight $(i,j)$ to facility $i$ and stop raising $\alpha_j$; all such clients $j$ are served by facility $i$;
- upon $\alpha_j = d_{ij}$ for a tentatively open facility $i$: connect client $j$ to facility $i$ and stop raising $\alpha_j$; such client $j$ is served by facility $i$;

**Phase I:**

**Phase II:**

construct graph $G(V,E)$ where $V=$\{tentatively open facilities\} and $(i_1, i_2) \in E$ if $\beta_{i_1,j}, \beta_{i_2,j} > 0$ for some client $j$;

find a maximal independent set $I$ of $G$ and permanently open facilities in $I$;

- directly connect every facility $i \in I$ to all clients $j$ that $i$ serves or $\beta_{ij} > 0$;
- for every remaining client $j$: connect $j$ to the nearest open facility; call such client indirectly connected;
Phase I:
Initially $\alpha = 0$, $\beta = 0$, no facility is open, no client is served; raise $\alpha_j$ for all client $j$ simultaneously at a **uniform continuous** rate:
- upon $\alpha_j = d_{ij}$ for a closed facility $i$: edge $(i,j)$ is **tight**; fix $\beta_{ij} = \alpha_j - d_{ij}$ as $\alpha_j$ being raised;
- upon $\sum_{j \in C} \beta_{ij} = f_i$: **tentatively open** facility $i$; connect all unserved clients $j$ with **tight** $(i,j)$ to facility $i$ and stop raising $\alpha_j$; all such clients $j$ are served by facility $i$;
- upon $\alpha_j = d_{ij}$ for a **tentatively open** facility $i$: connect client $j$ to facility $i$ and stop raising $\alpha_j$; such client $j$ is served by facility $i$;

Phase II:
construct graph $G(V,E)$ where $V=$ {tentatively open facilities} and $(i_1, i_2) \in E$ if $\beta_{i_1 j}, \beta_{i_2 j} > 0$ for some client $j$;
find a maximal independent set $I$ of $G$ and **permanently open** facilities in $I$;
- directly connect every facility $i \in I$ to all clients $j$ that $i$ serves or $\beta_{ij} > 0$;
- for every remaining client $j$: connect $j$ to the nearest open facility; call such client **indirectly connected**;

$I$ is independent set

Every client is connected to **exact one** open facilities.

(feasible)
Primal:

\[
\text{min } \sum_{i \in F, j \in C} d_{ij} x_{ij} + \sum_{i \in F} f_i y_i \\
\text{s.t. } y_i - x_{ij} \geq 0, \quad \forall i \in F, j \in C \\
\sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C \\
x_{ij}, y_i \in \{0, 1\}, \quad \forall i \in F, j \in C
\]

Dual-relax:

\[
\text{max } \sum_{j \in C} \alpha_j \\
\text{s.t. } \alpha_j - \beta_{ij} \leq d_{ij}, \quad \forall i \in F, j \in C \\
\sum_{j \in C} \beta_{ij} \leq f_i, \quad \forall i \in F \\
\alpha_j, \beta_{ij} \geq 0, \quad \forall i \in F, j \in C
\]

\(\alpha_j\): amount of value paid by client \(j\) to all facilities

\(\beta_{ij} \geq \alpha_j - d_{ij}\): payment to facility \(i\) by client \(j\) (after deduction)

Complimentary slackness conditions:

- if ideally held
- \(x_{ij} = 1 \Rightarrow \alpha_j - \beta_{ij} = d_{ij}\)
- \(\alpha_j > 0 \Rightarrow \sum_{i \in F} x_{ij} = 1\)
- \(\sum_{j \in C} \beta_{ij} = f_i\)
- \(\beta_{ij} > 0 \Rightarrow y_i = x_{ij}\)
Phase I:
Initially \( \alpha = 0, \beta = 0 \), no facility is open, no client is served; raise \( \alpha_j \) for all client \( j \) simultaneously at a uniform continuous rate:
- upon \( \alpha_j = d_{ij} \) for a closed facility \( i \): edge \((i,j)\) is tight; fix \( \beta_{ij} = \alpha_j - d_{ij} \) as \( \alpha_j \) being raised;
- upon \( \sum_{j \in C} \beta_{ij} = f_i \): tentatively open facility \( i \); connect all unserved clients \( j \) with tight \((i,j)\) to facility \( i \) and stop raising \( \alpha_j \); all such clients \( j \) are served by facility \( i \);
- upon \( \alpha_j = d_{ij} \) for a tentatively open facility \( i \): connect client \( j \) to facility \( i \) and stop raising \( \alpha_j \); such client \( j \) is served by facility \( i \);

Phase II:
construct graph \( G(V,E) \) where \( V=\{\)tentatively open facilities\(\} \) and \( (i_1, i_2) \in E \) if \( \beta_{i_1 j}, \beta_{i_2 j} > 0 \) for some client \( j \);
find a maximal independent set \( I \) of \( G \) and permanently open facilities in \( I \):
- directly connect every facility \( i \in I \) to all clients \( j \) that \( i \) serves or \( \beta_{ij} > 0 \);
- for every remaining client \( j \): connect \( j \) to the nearest open facility; call such client indirectly connected;

\[
\text{SOL} = \sum_{i \in I} f_i + \sum_{j: \text{directly connected}} d_{\phi(j),j} + \sum_{j: \text{indirectly connected}} d_{\phi(j),j} \leq 3 \sum_{j \in C} \alpha_j \leq 3 \text{OPT}
\]

\( \phi(j) = \begin{cases} 
\text{some } i \text{ that } \beta_{ij} = \alpha_j - d_{ij} & \text{if } j \text{ directly connected} \\
\text{nearest facility in } I & \text{if } j \text{ indirectly connected}
\end{cases} \)
Phase I:
Initially \( \alpha = 0, \beta = 0 \), no facility is open, no client is served;
raise \( \alpha_j \) for all client \( j \) simultaneously at a *uniform continuous* rate:
- upon \( \alpha_j = d_{ij} \) for a closed facility \( i \): edge \((i,j)\) is *tight*; fix \( \beta_{ij} = \alpha_j - d_{ij} \) as \( \alpha_j \) being raised;
- upon \( \sum_{j \in C} \beta_{ij} = f_i \): tentatively open facility \( i \); connect all *unserved* clients \( j \) with *tight* \((i,j)\) to facility \( i \) and stop raising \( \alpha_j \); all such clients \( j \) are *served* by facility \( i \);
- upon \( \alpha_j = d_{ij} \) for a tentatively open facility \( i \): connect client \( j \) to facility \( i \) and stop raising \( \alpha_j \); such client \( j \) is *served* by facility \( i \);

Phase II:
construct graph \( G(V,E) \) where \( V=\{\text{tentatively open facilities}\} \) and \((i_1,i_2) \in E \) if \( \beta_{i_1j}, \beta_{i_2j} > 0 \) for some client \( j \);
find a maximal independent set \( I \) of \( G \) and permanently open facilities in \( I \);
- directly connect every facility \( i \in I \) to all clients \( j \) that \( i \) serves or \( \beta_{ij} > 0 \);
- for every remaining client \( j \): connect \( j \) to the nearest open facility; call such client *indirectly connected*;

\[
\text{SOL} \leq 3 \text{ OPT}
\]

can be implemented *discretely*: in \( O(m \log m) \) time, \( m = |F||C| \)
- sort all edges \((i,j) \in F \times C\) by non-decreasing \( d_{ij} \)
- dynamically maintain the time of next event by heap
**Instance:** set $F$ of facilities; set $C$ of clients; facility opening costs $f: F \rightarrow [0, \infty)$; connection metric $d: F \times C \rightarrow [0, \infty)$; Find $\phi: C \rightarrow I \subseteq F$ to minimize $\sum_{j \in C} d_{\phi(j), j} + \sum_{i \in I} f_i$

$$\min \sum_{i \in F, j \in C} d_{ij} x_{ij} + \sum_{i \in F} f_i y_i$$

**s.t.**

$y_i - x_{ij} \geq 0, \quad \forall i \in F, j \in C$

$\sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C$

$x_{ij}, y_i \in \{0, 1\}, \quad \forall i \in F, j \in C$

- Integrality gap = 3
- no poly-time <1.463-approx. algorithm unless $\text{NP}=\text{P}$
- [Li 2011] 1.488-approx. algorithm
**k-Median**

**Instance:** set $F$ of facilities; set $C$ of clients; connection metric $d: F \times C \to [0, \infty)$;

Find a subset $I \subseteq F$ of $\leq k$ opening facilities and a way $\phi: C \to I$ of connecting all clients to them such that the total cost $\sum_{j \in C} d_{\phi(j), j}$ is minimized.