Piecewise Linear Approximation of Bi-level Problems in Transportation

Ximing Wang
Dr. Siriphong Lawphongpanich
Dr. Yafeng Yin

University of Florida
October 8, 2013

Introduction

- Toll design problem is to determine optimal locations of tolling facilities and toll prices when there are construction costs and a budget for constructing the facilities.

- This problem can be formulated as a bi-level optimization problem.
  - The decisions in the upper level determine where to toll and how much to charge
  - The decisions in the lower level correspond to travelers choosing routes with the least generalized cost (time plus tolls) to reach their destinations

- In literature, the bi-level problem has been approximated using a mixed-integer program where the objective and constraints are linear.
  - The equilibrium conditions are based on variational inequalities (VI)
  - Need to obtain extreme points and the corresponding inequality constraints contain bilinear terms that are neither convex nor concave

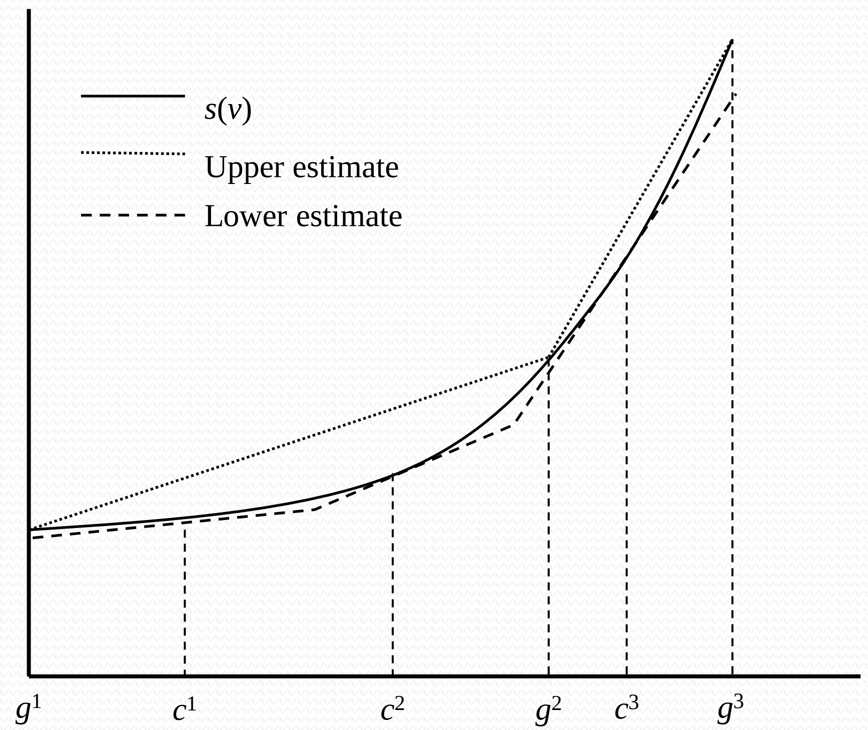
Introduction

- We use two piecewise linear functions to approximate the nonlinear functions in the problem, all of which are convex.
  - For each nonlinear function, one piecewise-linear function overestimates it and the other underestimates instead.
  - These piecewise linear functions do not require any binary variable to implement under mild conditions.
- We ensure user equilibrium via the KKT conditions in terms of link flows.
  - This makes the generation of paths or extreme points unnecessary.
- Under mild conditions, the algorithm either produces an optimal solution to the original problem after a finite number of iterations or generates a sequence of solutions that converges to an optimal one in the limit.

Congestion Pricing (CP) Problem

$$
\begin{array}{llll}
\min\limits_{\beta,u,v,x,\rho,\sigma,w} & \sum_{(i,j)\in\mathcal{A}} f_{ij}(v_{ij}) & & \\
\text{s.t.} & 0 \le \beta_{ij} \le \beta_{max} u_{ij} & \forall (i,j)\in\mathcal{A} & (1) \\
& \sum_{(i,j)\in\mathcal{A}} c_{ij} u_{ij} \le b & & (2) \\
& v_{ij} = \sum_{k\in K} x_{ij}^k & \forall (i,j)\in\mathcal{A} & (3) \\
& Ax^k = d_k E_k, & \forall k\in\mathcal{K} & (4) \\
& s_{ij}(v_{ij}) + \beta_{ij} - (\rho_i^k - \rho_j^k) = \sigma_{ij}^k, & \forall (i,j)\in\mathcal{A}, k\in\mathcal{K} & (5) \\
& x_{ij}^k \le d_k w_{ij}^k, & \forall (i,j)\in\mathcal{A}, k\in\mathcal{K} & (6) \\
& \sigma_{ij}^k \le M(1 - w_{ij}^k) & \forall (i,j)\in\mathcal{A}, k\in\mathcal{K} & (7) \\
& \sigma_{ij}^k \ge 0, & \forall (i,j)\in\mathcal{A}, k\in\mathcal{K} & (8) \\
& x_{ij}^k \ge 0, & \forall (i,j)\in\mathcal{A}, k\in\mathcal{K} & (9) \\
& u_{ij} \in \{0,1\} & \forall (i,j)\in\mathcal{A} & (10) \\
& w_{ij}^k \in \{0,1\} & \forall (i,j)\in\mathcal{A}, k\in\mathcal{K} & (11) \\
& -M \le \rho_i^k \le 0 & \forall i\in\mathcal{N}, k\in\mathcal{K} & (12)
\end{array}
$$

- Link travel time $s_{ij}(v_{ij})$ and aggregate link delay $f_{ij}(v_{ij}) = v_{ij} s_{ij}(v_{ij})$
  - $s_{ij}(v_{ij}) > 0$, $s'_{ij}(v_{ij}) > 0$, $s''_{ij}(v_{ij}) \ge 0$
  - $f_{ij}(v_{ij})$ is strictly convex



- Lower estimate
  - Cut-defining points $c_{ij}^{1}, \cdots, c_{ij}^{n_{ij}^{c}}$
  - $s_{ij}\left(v_{ij}\right) \geq \max_{p \in \left\{1,\cdots,n_{ij}^{c}\right\}} \left\{s_{ij}\left(c_{ij}^{p}\right) + s_{ij}'\left(c_{ij}^{p}\right)\left(v_{ij} - c_{ij}^{p}\right)\right\}$ by convexity
- Upper estimate
  - Grid points $0 = g_{ij}^{1} < g_{ij}^{2} < \cdots < g_{ij}^{n_{ij}^{g}} = d_{\mathcal{N}}$
  - $s_{ij}\left(v_{ij}\right) \leq \sum_{p=1}^{n_{ij}^{g}} \lambda_{ij}^{p} s_{ij}(g_{ij}^{p})$ for convex combination of grid points

CP Approximation

- Objective function $f_{ij}(v_{ij})$
  - Replaced with the average of the lower and upper estimates
  - $\frac{1}{2}\left(\max_{p=1,\cdots,n_{ij}^{c}}\left\{f_{ij}\left(c_{ij}^{p}\right)+f'_{ij}(c_{ij}^{p})(v_{ij}-c_{ij}^{p})\right\}+\sum_{p=1}^{n_{ij}^{g}}\lambda_{ij}^{p}f_{ij}\left(g_{ij}^{p}\right)\right)$
- Link travel time $s_{ij}(v_{ij})$
  - Replaced with auxiliary variable y_{ij} between the lower and upper estimates
  - $\max_{p=1,\cdots,n_{ij}^{c}}\left\{s_{ij}\left(c_{ij}^{p}\right)+s'_{ij}(c_{ij}^{p})(v_{ij}-c_{ij}^{p})\right\}\le y_{ij}\le\sum_{p=1}^{n_{ij}^{g}}\lambda_{ij}^{p}s_{ij}\left(g_{ij}^{p}\right)$

ACP problem

$$
\begin{aligned}
\min_{\beta,u,v,x,\rho,\sigma,w,z,\lambda,y} \quad & \frac{1}{2}\left(\sum_{p=1}^{n_{ij}^{c}} z_{ij} + \sum_{p=1}^{n_{ij}^{g}} \lambda_{ij}^{p} f_{ij}\left(g_{ij}^{p}\right)\right) && \\
\text{s.t.} \quad & f_{ij}\left(c_{ij}^{p}\right) + f_{ij}'\left(c_{ij}^{p}\right)\left(v_{ij} - c_{ij}^{p}\right) \le z_{ij}, & \forall p = 1,\cdots,n_{ij}^{c}, (i,j) \in \mathcal{A} & \quad (13)\\
& 0 \le \beta_{ij} \le \beta_{max} u_{ij} & \forall (i,j) \in \mathcal{A} & \\
& \sum_{(i,j)\in\mathcal{A}} c_{ij} u_{ij} \le b & & \\
& v_{ij} = \sum_{k\in K} x_{ij}^{k} & \forall (i,j) \in \mathcal{A} & \\
& v_{ij} = \sum_{p=1}^{n_{ij}^{g}} \lambda_{ij}^{p} g_{ij}^{p} & \forall (i,j) \in \mathcal{A} & \quad (14)\\
& A x^{k} = d_{k} E_{k}, & \forall k \in \mathcal{K} & \\
& y_{ij} + \beta_{ij} - (\rho_{i}^{k} - \rho_{j}^{k}) = \sigma_{ij}^{k}, & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& s_{ij}\left(c_{ij}^{p}\right) + s_{ij}'\left(c_{ij}^{p}\right)\left(v_{ij} - c_{ij}^{p}\right) \le y_{ij}, & \forall p = 1,\cdots,n_{ij}^{c}, (i,j) \in \mathcal{A} & \quad (15)\\
& y_{ij} \le \sum_{p=1}^{n_{ij}^{g}} \lambda_{ij}^{p} s_{ij}\left(g_{ij}^{p}\right) & \forall (i,j) \in \mathcal{A} & \quad (16)\\
& x_{ij}^{k} \le d_{k} w_{ij}^{k}, & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& \sigma_{ij}^{k} \le M(1 - w_{ij}^{k}) & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& \sigma_{ij}^{k} \ge 0, & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& x_{ij}^{k} \ge 0, & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& u_{ij} \in \{0,1\} & \forall (i,j) \in \mathcal{A} & \\
& w_{ij}^{k} \in \{0,1\} & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& -M \le \rho_{i}^{k} \le 0 & \forall i \in \mathcal{N}, k \in \mathcal{K} & \\
& \sum_{p=1}^{n_{ij}^{g}} \lambda_{ij}^{p} = 1, & \forall (i,j) \in \mathcal{A} & \quad (17)\\
& \lambda_{ij}^{p} \ge 0, & \forall p = 1,\cdots,n_{ij}^{g}, (i,j) \in \mathcal{A} & \quad (18)
\end{aligned}
$$

New cut-point defining

- ACP problem is linear and contains binary variables
  - Let $\hat{\pi} = (\hat{\beta}, \hat{u}, \hat{v}, \hat{x}, \hat{\rho}, \hat{\sigma}, \hat{w}, \hat{z}, \hat{\lambda}, \hat{y})$ be an optimal solution to the ACP

- The optimal link flows can be used as a new cut-defining points
  - Set $c_{ij}^{n_{ij}^{c}+1} = \hat{v}_{ij}$

New grid-point finding

- When the binary variables u_{ij} and w_{ij}^{k} are set to $\hat{u}_{ij}$ and $\hat{w}_{ij}^{k}$, ACP reduces to LP with $\left(\hat{\beta}, \hat{v}, \hat{x}, \hat{\rho}, \hat{\sigma}, \hat{z}, \hat{\lambda}, \hat{y}\right)$ as an optimal solution
  - The dual problem associated with this linear program also has a finite optimal solution
  - Let $\theta_{ij}, \psi_{ij},$ and ζ_{ij} be an optimal dual variable associated with the constraints in eq. (14), (16), and (17)
  - Because these dual values are optimal, the reduced costs of the current set of grid points must be nonnegative
    - $f_{ij}\left(g_{ij}^{p}\right) - \theta_{ij} g_{ij}^{p} - \psi_{ij} s_{ij}\left(g_{ij}^{p}\right) + \zeta_{ij} \geq 0$

New grid-point finding

$$\min_{\beta,u,v,x,\rho,\sigma,w,z,\lambda,y} \quad \frac{1}{2}\left(\sum_{p=1}^{n_{ij}^c} z_{ij} + \sum_{p=1}^{n_{ij}^g} \lambda_{ij}^p f_{ij}\left(g_{ij}^p\right)\right)$$

$$\begin{array}{llll}
\text{s.t.} & f_{ij}\left(c_{ij}^p\right) + f_{ij}'\left(c_{ij}^p\right)\left(v_{ij} - c_{ij}^p\right) \le z_{ij}, & \forall p = 1,\cdots,n_{ij}^c, (i,j) \in \mathcal{A} & (13) \\
& 0 \le \beta_{ij} \le \beta_{max} u_{ij} & \forall (i,j) \in \mathcal{A} & \\
& \sum_{(i,j)\in\mathcal{A}} c_{ij} u_{ij} \le b & & \\
& v_{ij} = \sum_{k\in K} x_{ij}^k & \forall (i,j) \in \mathcal{A} & \\
& v_{ij} = \sum_{p=1}^{n_{ij}^g} \lambda_{ij}^p g_{ij}^p & \forall (i,j) \in \mathcal{A} & (14) \quad \theta_{ij} \\
& A x^k = d_k E_k, & \forall k \in \mathcal{K} & \\
& y_{ij} + \beta_{ij} - (\rho_i^k - \rho_j^k) = \sigma_{ij}^k, & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& s_{ij}\left(c_{ij}^p\right) + s_{ij}'\left(c_{ij}^p\right)\left(v_{ij} - c_{ij}^p\right) \le y_{ij}, & \forall p = 1,\cdots,n_{ij}^c, (i,j) \in \mathcal{A} & (15) \\
& y_{ij} \le \sum_{p=1}^{n_{ij}^g} \lambda_{ij}^p s_{ij}\left(g_{ij}^p\right) & \forall (i,j) \in \mathcal{A} & (16) \quad \psi_{ij} \\
& x_{ij}^k \le d_k w_{ij}^k, & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& \sigma_{ij}^k \le M(1 - w_{ij}^k) & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& \sigma_{ij}^k \ge 0, & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& x_{ij}^k \ge 0, & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& u_{ij} \in \{0,1\} & \forall (i,j) \in \mathcal{A} & \\
& w_{ij}^k \in \{0,1\} & \forall (i,j) \in \mathcal{A}, k \in \mathcal{K} & \\
& -M \le \rho_i^k \le 0 & \forall i \in \mathcal{N}, k \in \mathcal{K} & \\
& \sum_{p=1}^{n_{ij}^g} \lambda_{ij}^p = 1, & \forall (i,j) \in \mathcal{A} & (17) \quad \zeta_{ij} \\
& \lambda_{ij}^p \ge 0, & \forall p = 1,\cdots,n_{ij}^g, (i,j) \in \mathcal{A} & (18)
\end{array}$$

New grid-point finding

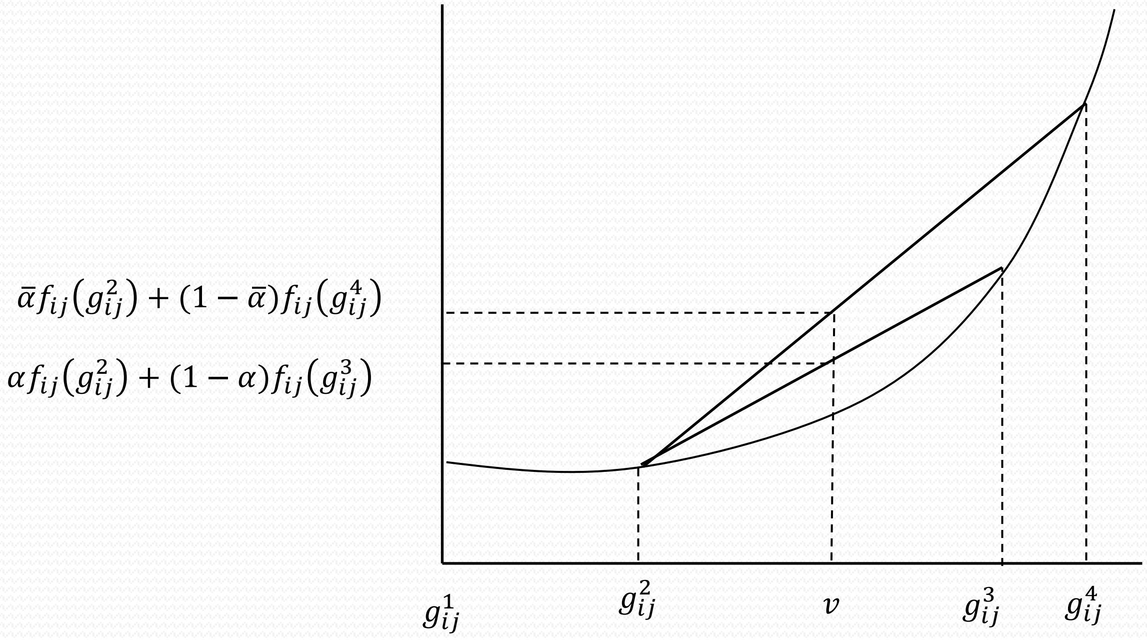
- Assume that a new grid point $g_{ij}^{n_{ij}^g+1} \neq g_{ij}^p, \forall p \in \{1, \cdots, n_{ij}^g\}$ is given, the reduced cost associate with new grid point is $r_{ij}\left(g_{ij}^{n_{ij}^g+1}\right) = f_{ij}\left(g_{ij}^{n_{ij}^g+1}\right) - \theta_{ij} g_{ij}^{n_{ij}^g+1} - \psi_{ij} s_{ij}\left(g_{ij}^{n_{ij}^g+1}\right) + \zeta_{ij}$.

- If $r_{ij}\left(g_{ij}^{n_{ij}^g+1}\right) < 0$, then adding $g_{ij}^{n_{ij}^g+1}$ to the ACP problem and pivoting or making $\lambda_{ij}^{n_{ij}^g+1}$ basic must reduce the upper estimate of $f_{ij}(\cdot)$ in the objective of the ACP problem.

- **Grid-point finding problem**
  - $g_{ij}^{n_{ij}^g+1} = \underset{0 \leq g_{ij} \leq d_{\mathcal{N}}}{\operatorname{argmin}} \{f_{ij}(g_{ij}) - \theta_{ij} g_{ij} - \psi_{ij} s_{ij}(g_{ij}) + \zeta_{ij}\}$
  - If the optimal objective value of the above problem is negative, adding the new grid point would improve the upper estimate of $f_{ij}(\cdot)$

Properties of ACP

- v_{ij} is a convex combination of at most two adjacent grid points



Properties of ACP

$h(\cdot)$ is strongly quasi-convex if the following holds for all $x^1 \neq x^2$

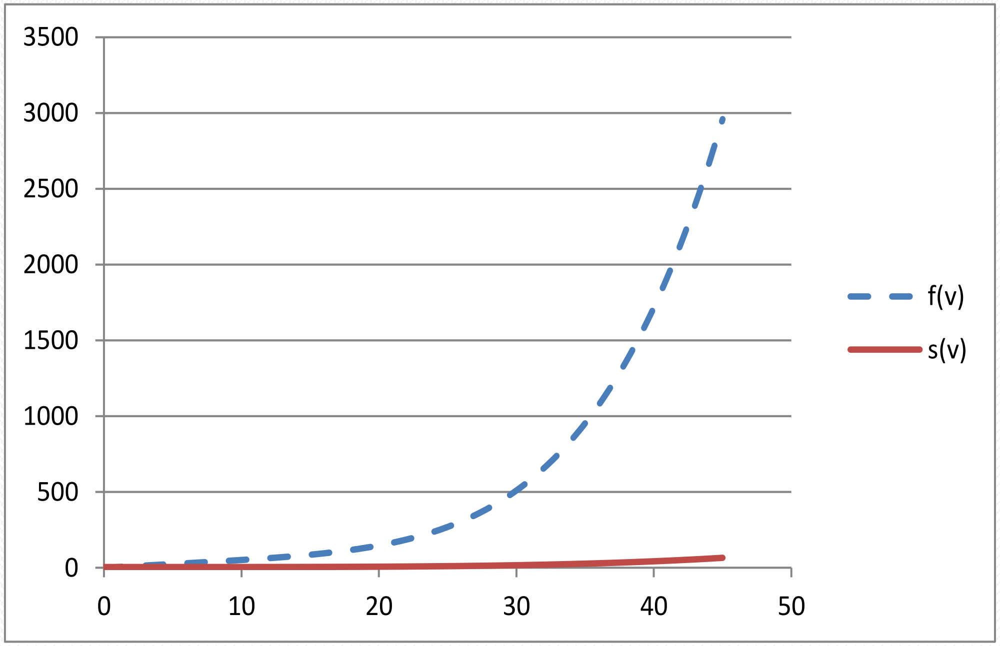
$$h(\alpha x^1 + (1-\alpha)x^2) < \max\{h(x^1), h(x^2)\}, \forall \alpha \in (0,1).$$

Theorem: If $r_{ij}(g_{ij}) = f_{ij}(g_{ij}) - \theta_{ij} g_{ij} - \psi_{ij} s_{ij}(g_{ij}) + \zeta_{ij}$ is strongly quasi-convex, then $\hat{v}_{ij}$ must be a convex combination of at most two adjacent grid points.

Theorem: Let $r_{ij}(g_{ij}) = f_{ij}(g_{ij}) - \theta_{ij} g_{ij} - \psi_{ij} s_{ij}(g_{ij}) + \zeta_{ij}$ be strongly quasi-convex, if $\hat{v}_{ij} = \alpha g_{ij}^m + (1-\alpha) g_{ij}^{m+1}$ for some $\alpha \in (0,1)$ and $m \in \{1, \cdots, n_{ij}^g\}$, then the optimal solution $g_{ij}^{n_{ij}^g+1}$ must be in the interval (g^m, g^{m+1}).

Properties of ACP

- $r_{ij}(g_{ij}) = f_{ij}(g_{ij}) - \theta_{ij} g_{ij} - \psi_{ij} s_{ij}(g_{ij}) + \zeta_{ij}$
- Relative to $f_{ij}(\cdot)$, $s_{ij}(\cdot)$ appears approximately linear
- When $s_{ij}(\cdot)$ is linear, the reduced cost function $r_{ij}(\cdot)$ is strictly convex



Cutting-Plane and Grid-Point (CPGP) Algorithm

Step 0: Let $g_{ij}^p, \forall p = 1, \cdots, n_{ij}^g$ and $c_{ij}^p, \forall p = 1, \cdots, n_{ij}^c$, be grid and cut-defining points for arc (i, j). Set $\tau = 1$.

Step 1: Let $\big(\beta(\tau), u(\tau), v(\tau), x(\tau), \rho(\tau), \sigma(\tau), w(\tau), z(\tau), \lambda(\tau), y(\tau)\big)$ solve the ACP problem.

Step 2: For each $(i, j) \in \mathcal{A}$, let

$$g_{ij}(\tau) = \underset{0 \le g_{ij} \le d_{\mathcal{N}}}{\operatorname{argmin}} \{ f_{ij}(g_{ij}) - \theta_{ij} g_{ij} - \psi_{ij} s_{ij}(g_{ij}) + \zeta_{ij} \}$$

where (θ, ψ, ζ) denotes optimal dual variables associated with eq. (14), (16), and (17) in a linear program obtained by setting $u_{ij} = u_{ij}(\tau)$ and $w = w(\tau)$ in the ACP problem.

Set $r_{ij}(g_{ij}(\tau)) = f_{ij}\left(g_{ij}(\tau)\right) - \theta_{ij} g_{ij}(\tau) - \psi_{ij} s_{ij}\left(g_{ij}(\tau)\right) + \zeta_{ij}$.

Cutting-Plane and Grid-Point (CPGP) Algorithm

<u>Step 3</u>: For each $(i,j) \in \mathcal{A}$, calculate gaps in the lower estimates as follows:

$$\Delta_{ij}^{f} = f_{ij}(v(\tau)) - \max_{p \in \{1,\cdots,n_{ij}^{c}\}} \left\{ f_{ij}(c_{ij}^{p}) + f_{ij}'(c_{ij}^{p})\left(v(\tau) - c_{ij}^{p}\right) \right\}$$

$$\Delta_{ij}^{s} = s_{ij}(v(\tau)) - \max_{p \in \{1,\cdots,n_{ij}^{c}\}} \left\{ s_{ij}(c_{ij}^{p}) + s_{ij}'(c_{ij}^{p})\left(v(\tau) - c_{ij}^{p}\right) \right\}$$

If $\Delta_{ij}^{f} = 0, \Delta_{ij}^{s} = 0$, and $r_{ij}\left(g_{ij}(\tau)\right) \geq 0, \forall (i,j) \in \mathcal{A}$, then stop and $\left(\beta(\tau), u(\tau), v(\tau), x(\tau), \rho(\tau), \sigma(\tau), w(\tau)\right)$ is optimal to the CP problem. Otherwise, for each $(i,j) \in \mathcal{A}$, do the following and return to Step 1.

a) Set $c_{ij}^{n_{ij}^{c}+1} = v(\tau)$ and $n_{ij}^{c} = n_{ij}^{c} + 1$, if $\Delta_{ij}^{f} > 0$ or $\Delta_{ij}^{s} > 0$

b) Set $g_{ij}^{n_{ij}^{g}+1} = g_{ij}(\tau)$ and $n_{ij}^{g} = n_{ij}^{g} + 1$, if $r_{ij}\left(g_{ij}(\tau)\right) < 0$.

Properties of CPGP Algorithm

- When the grid-point finding problem in Step 2 has a unique solution and $r_{ij}(\tau) = 0$, the optimal aggregate link flow $v_{ij}(\tau)$ obtained in Step 1 equal to one of the grid points.
- When in Step 3, there is no gap and no grid point with a negative reduced cost, then the algorithm can stop because parts of the solution in Step 1 is optimal to the CP problem.

Convergence Analysis

- Let $\pi(\tau)$ denote the solution of the ACP problem in iteration τ
  - $\pi(\tau) = \big(\beta(\tau), u(\tau), v(\tau), x(\tau), \rho(\tau), \sigma(\tau), w(\tau), z(\tau), \lambda(\tau), y(\tau)\big)$

- Assume CPGP algorithm generates an infinite sequence $\{\pi(\tau)\}_\tau$
  - $u(\tau)$ and $w(\tau)$ are binary vectors with finite number of elements
  - There exist $\ddot{u}$ and $\ddot{w}$ such that $u(\tau) = \ddot{u}$ and $w(\tau) = \ddot{w}$ infinitely often
  - Subset $\Omega \subset \{1, 2, \cdots, \infty\}$ such that $u(\tau) = \ddot{u}$ and $w(\tau) = \ddot{w}, \forall \tau \in \Omega$
  - Setting $u_{ij} = \ddot{u}$ and $w_{ij}^k = \ddot{w}_{ij}^k$ renders the ACP problem to LP

- $\Omega_1 \subseteq \Omega$ such that $\{\pi(\tau)\}_{\tau \in \Omega_1}$ converges to $\ddot{\pi} = \big(\ddot{\beta}, \ddot{u}, \ddot{v}, \ddot{x}, \ddot{\rho}, \ddot{\sigma}, \ddot{w}, \ddot{z}, \ddot{\lambda}, \ddot{y}\big)$
  - For any Ω_1 that yields a convergent subsequence, $\ddot{\pi}$ solves the CP problem

Numerical Experiments

- The construction cost for all toll facilities is 1
  - The budget b is the maximum number of toll facilities to be constructed

- The initial number of grid points is 6
  - $g_{ij}^1 = 0, g_{ij}^2 = cap_{ij}, g_{ij}^3 = 2 \times cap_{ij}, g_{ij}^4 = 3 \times cap_{ij}, g_{ij}^5 = 4 \times cap_{ij}$
  - g_{ij}^6 depends on b
    - Obtain SO flow and compute the associated externalities
    - Links with b largest externalities have tolls equal to their externalities
    - g_{ij}^6 is the tolled UE link flow

- CPGP algorithm termination criteria
  - The gaps for all underestimates relatively to their function values (link travel time and aggregate delay) and all reduced costs relative to the network delay is less than 1%

Numerical Experiments

Results from nine-node network

	CPGP				Ekstrom et al. (2012)		
Facilities	**Approx.**	**Actual**	**Iter.**	**CPU (sec)**	**Actual**	**CPU (sec)**	**Scheme**
5	2252.14	2254.45	4	4	2253.92	6440	*l*9
3	2281.22	2281.93	4	13	2281.97	33034	*l*10
1	2364.91	2361.42	3	8	2361.22	249	*l*3

Results from Sioux Falls

	CPGP				Ekstrom et al. (2012)		
Facilities	**Approx.**	**Actual**	**Iter.**	**CPU (sec)**	**Actual**	**CPU (sec)**	**Scheme**
5	4312.02	4316.96	6	8196	4328.24	41216	*l*2
4	4336.22	4339.90	4	5222	4345.19	27930	*l*4
1	4418.52	4428.69	4	4001	4437.65	5488	*l*1

Conclusions

- This talk proposes a piecewise linear approximation scheme for solving bi-level problems in transportation.
- The scheme allows a bi-level problem to be solved approximately as a linear integer program.
- The approximate solution can be further refined by adding additional linear pieces and solving the expanded linear integer program starting from a previous solution.
- Under mild conditions, the algorithm either produces an optimal solution to the original problem after a finite number of iterations or generates a sequence of solutions that converges to an optimal one in the limit.
- Numerical results show the efficiency of the algorithm.

Thank you!