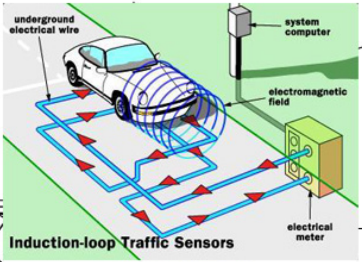
Estimation of game-theoretical systems with applications to urban transportation

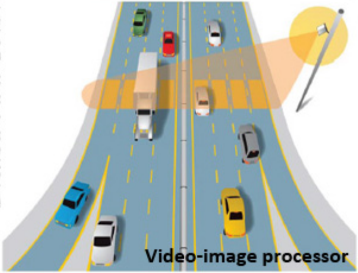
Jérôme Thai[1] Alexandre Bayen

[1]Department of Electrical Engineering & Computer Sciences
University of California at Berkeley

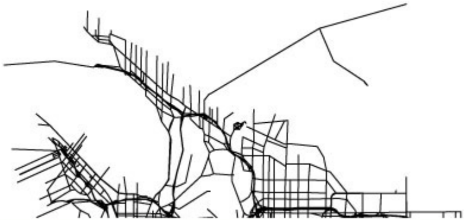
May 28, 2015

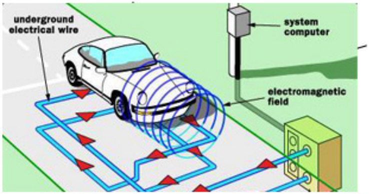
Limited sensing infrastructure





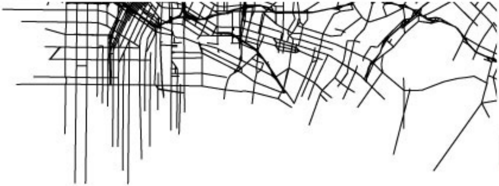
Limited sensing infrastructure

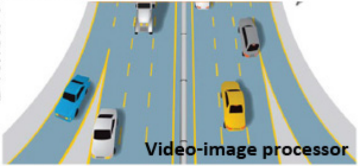




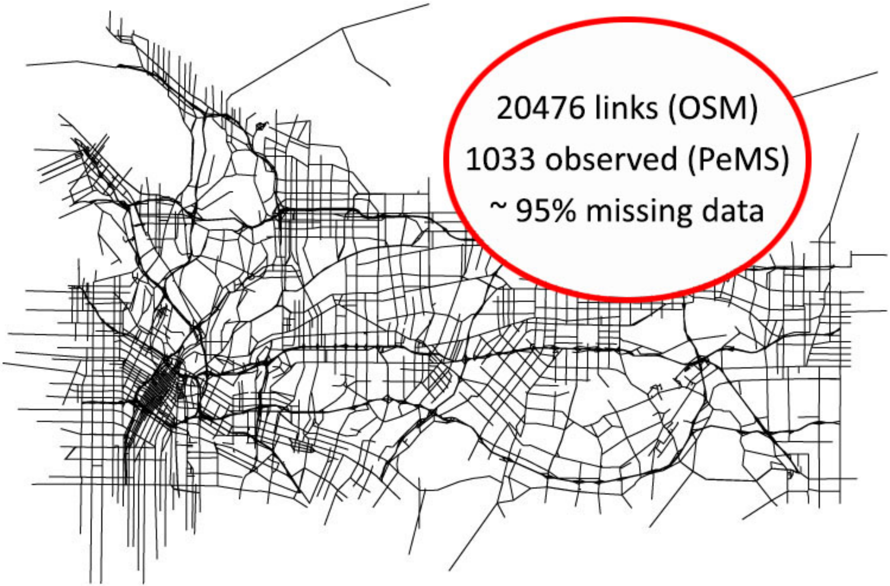
Sensor	Expected life	Cost/lane/year
Inductive loop detector	10 years	$746
Video image processor	10 years	$580

Middleton and Parker. *Initial Evaluation of Selected Detectors to Replace Inductive Loops on Freeways*, FHWA/TX-00/1439-7. Texas Transportation Institute, College Station, TX. April 2000.



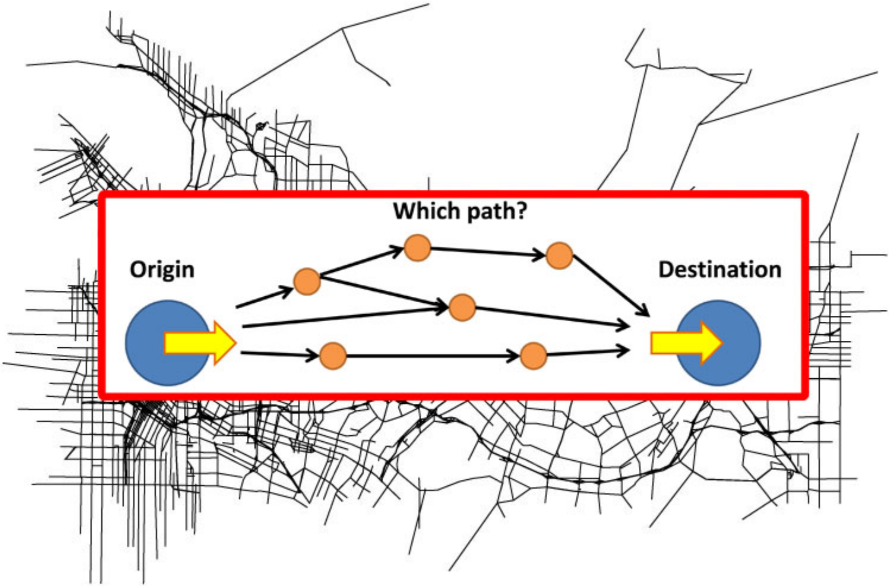


Sparsity of the data

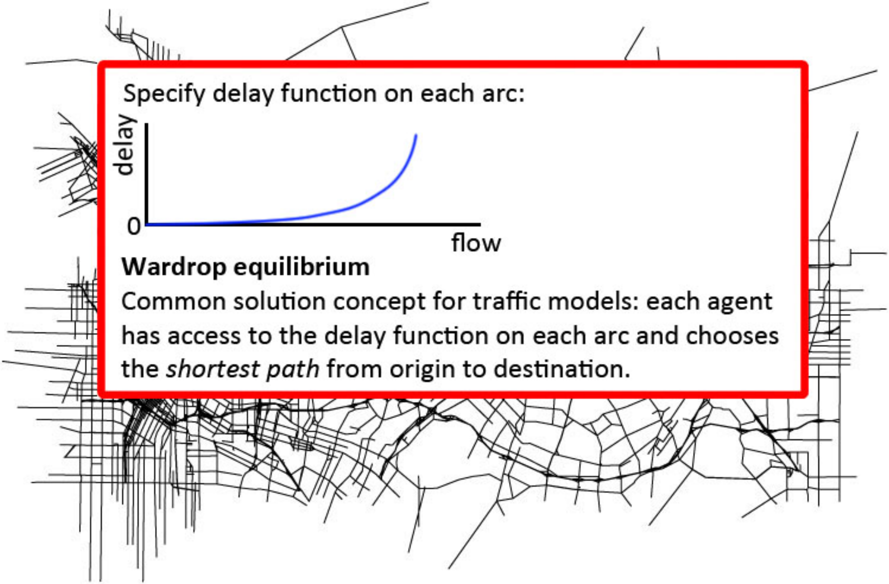


Quasi-static traffic assignment problem

Quasi-static traffic assignment problem



Quasi-static traffic assignment problem



Specify delay function on each arc:

Wardrop equilibrium

Common solution concept for traffic models: each agent has access to the delay function on each arc and chooses the *shortest path* from origin to destination.

Problem statement

We pose the inverse traffic assignment problem with missing data

Problem statement

We pose the inverse traffic assignment problem with missing data

- Traffic volumes resulting from rational behavior of agents on the road network are easily but **sparsely observable**.

Problem statement

We pose the inverse traffic assignment problem with missing data

- Traffic volumes resulting from rational behavior of agents on the road network are easily but **sparsely observable**.
- Delay functions are not directly observable.

Problem statement

We pose the inverse traffic assignment problem with missing data

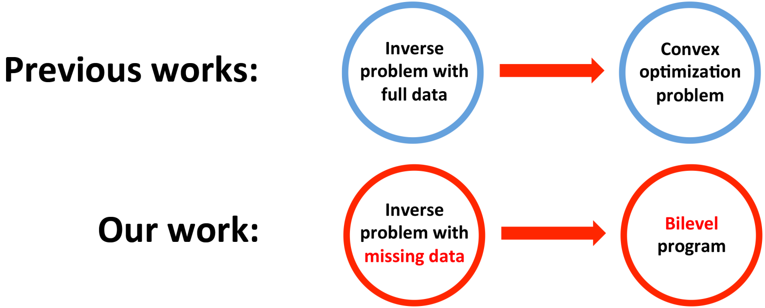
- Traffic volumes resulting from rational behavior of agents on the road network are easily but **sparsely observable**.
- Delay functions are not directly observable.
- How can we impute the delay functions from **partial observations** of equilibria?

Previous works assume full observations

- **Inverse convex optimization:** Keshavarz et al. (2011)
- **Inverse variational inequality:** Bertsimas et al. (2014)

Previous works assume full observations

- **Inverse convex optimization:** Keshavarz et al. (2011)
- **Inverse variational inequality:** Bertsimas et al. (2014)



Here we develop more complex tools combining ideas from:

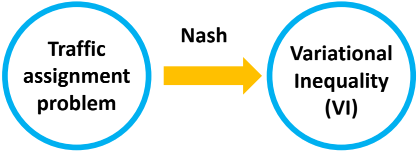
- Bilevel programming
- Computational mathematics
- Pareto optimization

Outline and contributions

Forward problem: **Traffic assignment problem**

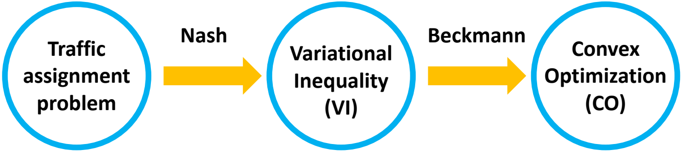
Outline and contributions

Forward problem:

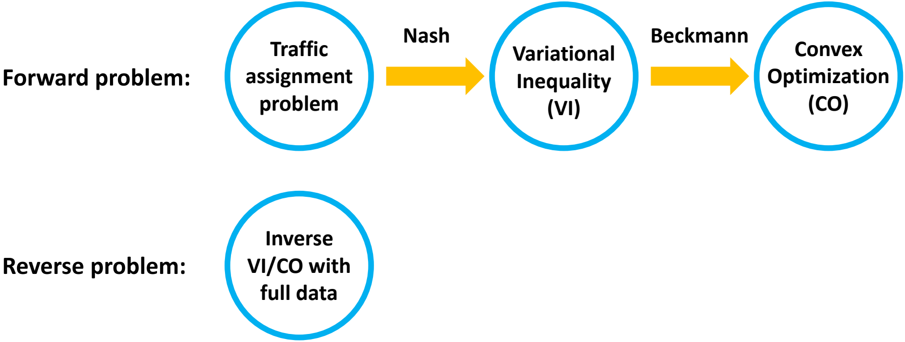


Outline and contributions

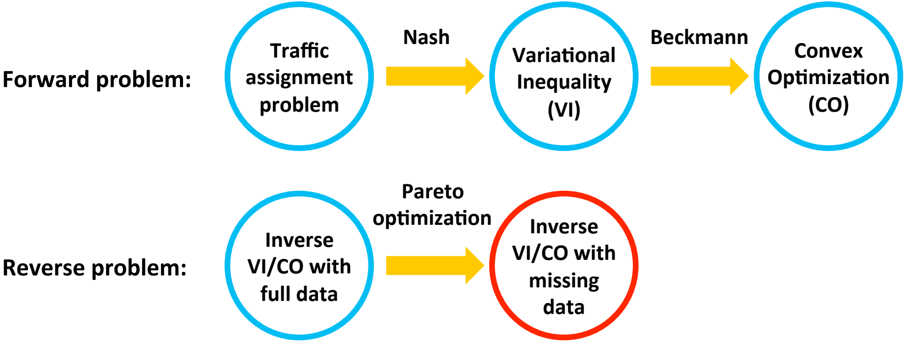
Forward problem:



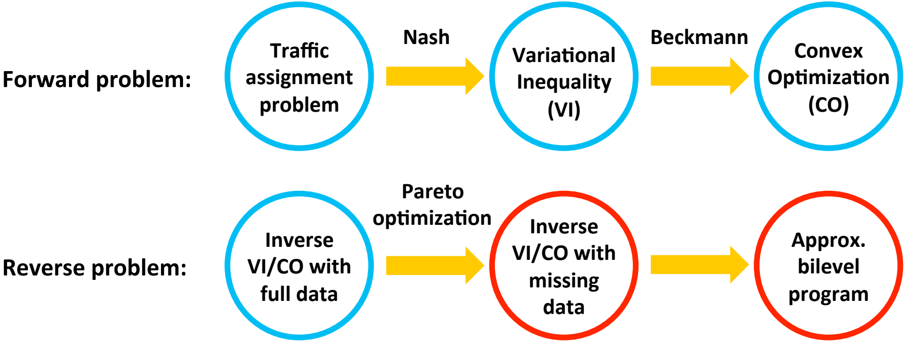
Outline and contributions



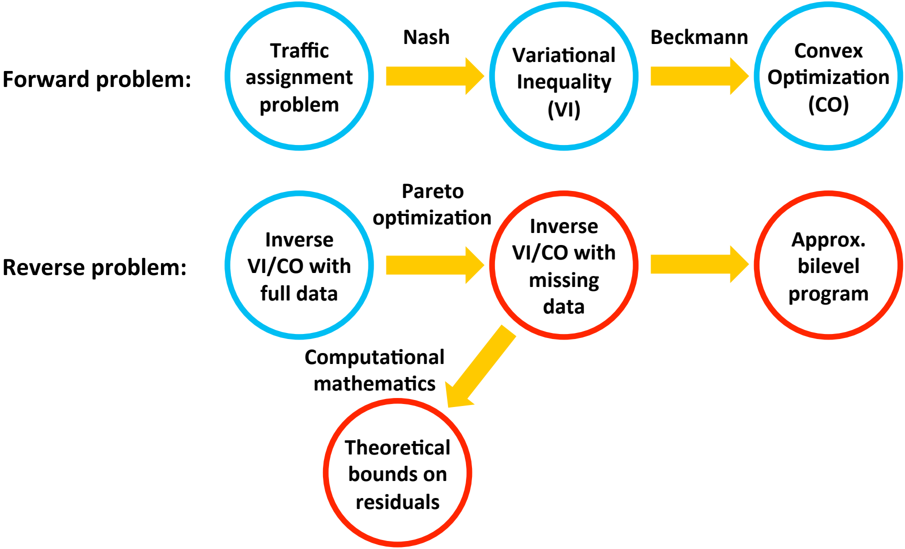
Outline and contributions



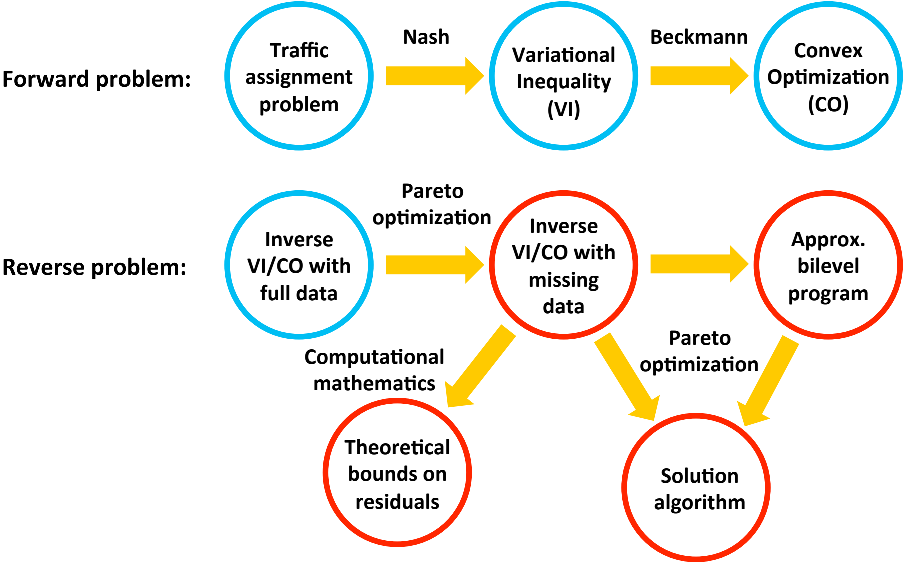
Outline and contributions



Outline and contributions



Outline and contributions



Outline

Inverse problem with missing data

Formulation as a Pareto optimization problem

Theoretical results and implementation

Optimization process and Variational inequality

Notations and assumptions:

- $\mathcal{K} := \{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0\}$ encodes the flow conservation.
- Arc delays are increasing, separable and encoded in map $F : \mathbb{R}^n \to \mathbb{R}^n$

Variational Inequality (VI) formulation

The flow vector $\mathbf{x}^\star \in \mathcal{K}$ is an eq. iff $F(\mathbf{x}^\star)^T(\mathbf{u} - \mathbf{x}^\star) \geq 0,\ \forall\, \mathbf{u} \in \mathcal{K}$.

Optimization process and Variational inequality

Notations and assumptions:

- $\mathcal{K} := \{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0\}$ encodes the flow conservation.
- Arc delays are increasing, separable and encoded in map $F : \mathbb{R}^n \to \mathbb{R}^n$

Variational Inequality (VI) formulation

The flow vector $\mathbf{x}^\star \in \mathcal{K}$ is an eq. iff $F(\mathbf{x}^\star)^T(\mathbf{u} - \mathbf{x}^\star) \geq 0,\ \forall\, \mathbf{u} \in \mathcal{K}$.

Beckmann: for the delay map F, $\exists\, f$ convex such that $F = \nabla f$

Theorem 1 (Beckmann et al. 1956)

The eq. is solution of a convex optimization program OP$(\mathcal{K}, f)$:
$\min\ f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0.$

Remarks:

- the potential f encodes the interaction between players.
- the VI is a first-order optimality condition.

Review of Inverse problem

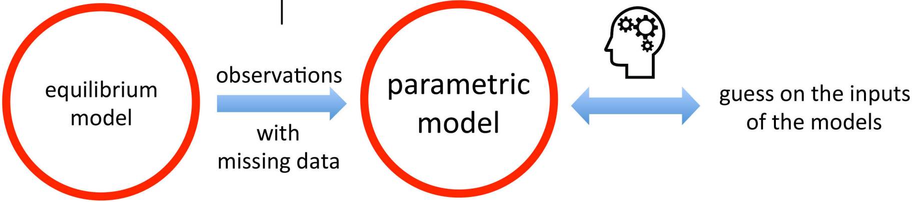
Equilibrium model

$(K, F(\cdot))$

Strategy set K
$K \subset \mathbb{R}^n$ closed convex

Payoffs function F
$F : \mathbb{R}^n \to \mathbb{R}^n$

Set of equilibria $\mathcal{S} \subset \mathbb{R}^n$
strategy vector $x \in \mathcal{S}$



Review of Inverse problem

Equilibrium model

$(K, F(\cdot))$

Strategy set K
$K \subset \mathbb{R}^n$ closed convex

Payoffs function F
$F : \mathbb{R}^n \to \mathbb{R}^n$

Set of equilibria $\mathcal{S} \subset \mathbb{R}^n$
strategy vector $x \in \mathcal{S}$

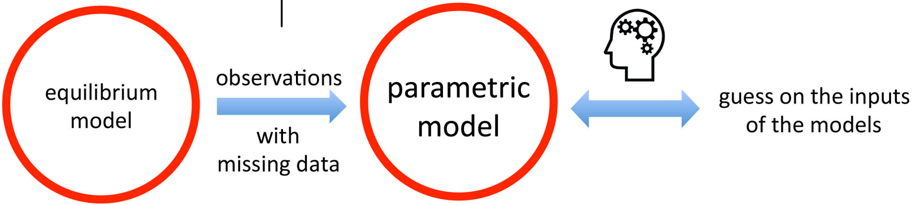
Parametric model

Assumes a parametric model
$(K, F(\cdot|\theta))$, $\theta \in \Theta$

Structural parameters $\theta \in \mathbb{R}^d$

Θ contains *enough* prior
information about the model

observes $z = Hx \in \mathbb{R}^p$
with $x \in \mathcal{S}$ and H the observer



Review of Inverse problem

Equilibrium model

$(K, F(\cdot))$

Strategy set K
$K \subset \mathbb{R}^n$ closed convex

Payoffs function F
$F : \mathbb{R}^n \to \mathbb{R}^n$

Set of equilibria $\mathcal{S} \subset \mathbb{R}^n$
strategy vector $x \in \mathcal{S}$

Parametric model

Assumes a parametric model $(K, F(\cdot|\theta))$, $\theta \in \Theta$

Structural parameters $\theta \in \mathbb{R}^d$

Θ contains *enough* prior information about the model

observes $z = Hx \in \mathbb{R}^p$
with $x \in \mathcal{S}$ and H the observer

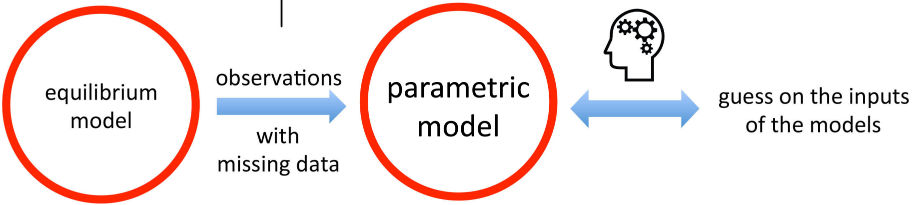
Mathematical program

$$\min_{x,\theta} \ \|Hx - z\| \quad \text{s.t.} \quad x \in \mathcal{S}(\theta), \ \theta \in \Theta$$

Find the best $\theta \in \Theta$ to

minimize the measurement residual $\|Hx - z\|$

such that x is an equilibrium of $F(K, F(\cdot|\theta))$,



Estimation of the highway network near Los Angeles

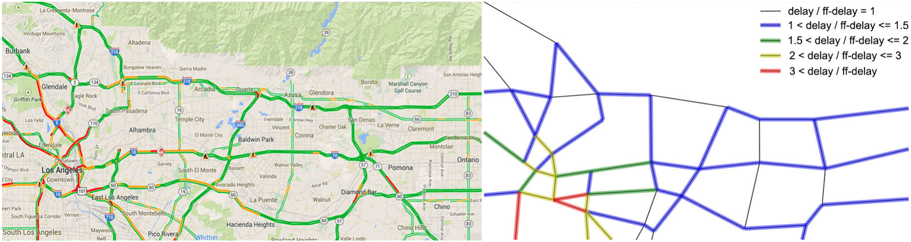


Figure : Highway network near Los Angeles

Estimation of the highway network near Los Angeles

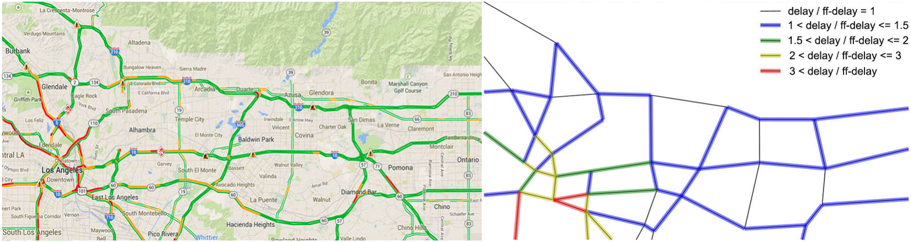


Figure : Highway network near Los Angeles

- True delay function: $s_a^{\text{true}}(v_a) = d_a \left(1 - \frac{3.5}{3} + \frac{3.5}{3 - v_a/m_a}\right)$
- Parametric delay: $s_a(v_a|\boldsymbol{\theta}) = d_a \left(1 + \sum_{i=1}^{6} \theta_i (v_a/m_a)^i\right)$

where $d_a =$ free flow delay, $m_a =$ number of lanes, $v_a =$ aggregate flow.

Estimation of the highway network near Los Angeles

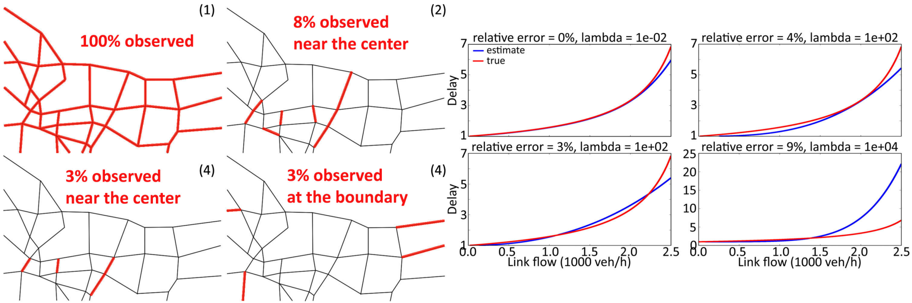


Figure : Delay function imputation.

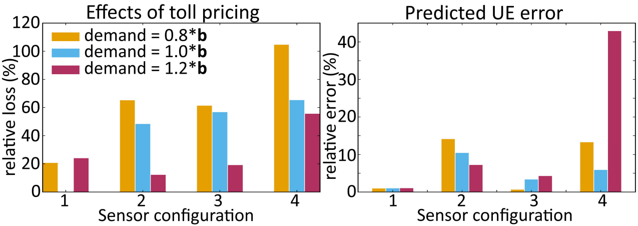


Figure : Toll pricing.

Outline

Inverse problem with missing data

Formulation as a Pareto optimization problem

Theoretical results and implementation

Primal-dual system and KKT conditions

Assumption: $\mathcal{K} = \{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b},\, \mathbf{x} \succeq 0\}$

Primal-dual system and KKT conditions

Assumption: $\mathcal{K} = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\, \mathbf{x} \succeq 0\}$

Theorem 3: primal-dual system (Facchinei 2003 Aghassi 2005)

$\mathbf{x}$ is solution to $\mathrm{VI}(\mathcal{K}, F)$ if and only if there exists $\mathbf{y}$ such that

$$
\begin{aligned}
&F(\mathbf{x})^T\mathbf{x} = \mathbf{b}^T\mathbf{y} \\
&\mathbf{A}\mathbf{x} = \mathbf{b},\, \mathbf{x} \succeq 0 \\
&\mathbf{A}^T\mathbf{y} \preceq F(\mathbf{x})
\end{aligned}
$$

Primal-dual system and KKT conditions

Assumption: $\mathcal{K} = \{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0\}$

Theorem 3: primal-dual system (Facchinei 2003 Aghassi 2005)

$\mathbf{x}$ is solution to VI($\mathcal{K}, F$) if and only if there exists $\mathbf{y}$ such that

$$F(\mathbf{x})^T\mathbf{x} = \mathbf{b}^T\mathbf{y}$$
$$\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0$$
$$\mathbf{A}^T\mathbf{y} \preceq F(\mathbf{x})$$

Theorem 4: KKT conditions (Harker 1989)

$\mathbf{x}$ is solution to VI($\mathcal{K}, F$) if and only if there exists $(\mathbf{y}, \boldsymbol{\pi})$ such that

$$F(\mathbf{x}) = \mathbf{A}^T\mathbf{y} + \boldsymbol{\pi}$$
$$\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0$$
$$\boldsymbol{\pi} \succeq 0,\ \mathbf{x}^T\boldsymbol{\pi} = 0$$

Note: If $F = \nabla f$, we can substitute VI($\mathcal{K}$,F) with OP($\mathcal{K}, f$).

Primal-dual system and KKT conditions

Assumption: $\mathcal{K} = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\, \mathbf{x} \succeq 0\}$

Theorem 3: primal-dual system (Facchinei 2003 Aghassi 2005)

$\mathbf{x}$ is solution to VI($\mathcal{K}, F$) if and only if there exists $\mathbf{y}$ such that

$$F(\mathbf{x})^T\mathbf{x} = \mathbf{b}^T\mathbf{y}$$
$$\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \quad \text{primal feasibility}$$
$$\mathbf{A}^T\mathbf{y} \preceq F(\mathbf{x}) \quad \text{dual feasibility}$$

Theorem 4: KKT conditions (Harker 1989)

$\mathbf{x}$ is solution to VI($\mathcal{K}, F$) if and only if there exists $(\mathbf{y}, \boldsymbol{\pi})$ such that

$$F(\mathbf{x}) = \mathbf{A}^T\mathbf{y} + \boldsymbol{\pi}$$
$$\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \quad \text{primal feasibility}$$
$$\boldsymbol{\pi} \succeq 0,\ \mathbf{x}^T\boldsymbol{\pi} = 0 \quad \text{dual feasibility}$$

Note: If $F = \nabla f$, we can substitute VI($\mathcal{K}$,F) with OP($\mathcal{K}, f$).

Primal-dual system and KKT conditions

Assumption: $\mathcal{K} = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0\}$

Theorem 3: primal-dual system (Facchinei 2003 Aghassi 2005)

$\mathbf{x}$ is solution to VI($\mathcal{K}, F$) if and only if there exists $\mathbf{y}$ such that

$$F(\mathbf{x})^T\mathbf{x} = \mathbf{b}^T\mathbf{y}$$
$$\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \quad \text{primal feasibility}$$
$$\mathbf{A}^T\mathbf{y} \preceq F(\mathbf{x}) \quad \text{dual feasibility}$$

Theorem 4: KKT conditions (Harker 1989)

$\mathbf{x}$ is solution to VI($\mathcal{K}, F$) if and only if there exists $(\mathbf{y}, \boldsymbol{\pi})$ such that

$$F(\mathbf{x}) = \mathbf{A}^T\mathbf{y} + \boldsymbol{\pi}$$
$$\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \quad \text{primal feasibility}$$
$$\boldsymbol{\pi} \succeq 0,\ \mathbf{x}^T\boldsymbol{\pi} = 0 \quad \text{dual feasibility}$$

Note: If $F = \nabla f$, we can substitute VI($\mathcal{K}$,F) with OP($\mathcal{K}, f$).

Residual functions

Definition: residual functions

Nonnegative functions r_{PD} and r_{KKT} such that

$$r_{\mathrm{PD}}(\mathbf{x}, \mathbf{y}) = 0 \iff (\mathbf{x}, \mathbf{y}) \text{ solution to primal-dual system}$$

$$r_{\mathrm{KKT}}(\mathbf{x}, \mathbf{y}, \boldsymbol{\pi}) = 0 \iff (\mathbf{x}, \mathbf{y}, \boldsymbol{\pi}) \text{ solution to KKT system}$$

Residual functions

Definition: residual functions

Nonnegative functions r_{PD} and r_{KKT} such that

$$r_{\text{PD}}(\mathbf{x}, \mathbf{y}) = 0 \iff (\mathbf{x}, \mathbf{y}) \text{ solution to primal-dual system}$$
$$r_{\text{KKT}}(\mathbf{x}, \mathbf{y}, \boldsymbol{\pi}) = 0 \iff (\mathbf{x}, \mathbf{y}, \boldsymbol{\pi}) \text{ solution to KKT system}$$

Classic residual associated to the primal-dual system

$$r_{\text{PD}}(\mathbf{x}, \mathbf{y}) = F(\mathbf{x})^T \mathbf{x} - \mathbf{b}^T \mathbf{y}$$

Residual functions

Definition: residual functions

Nonnegative functions r_{PD} and r_{KKT} such that

$$r_{\text{PD}}(\mathbf{x}, \mathbf{y}) = 0 \iff (\mathbf{x}, \mathbf{y}) \text{ solution to primal-dual system}$$
$$r_{\text{KKT}}(\mathbf{x}, \mathbf{y}, \boldsymbol{\pi}) = 0 \iff (\mathbf{x}, \mathbf{y}, \boldsymbol{\pi}) \text{ solution to KKT system}$$

Classic residual associated to the primal-dual system

$$r_{\text{PD}}(\mathbf{x}, \mathbf{y}) = F(\mathbf{x})^T \mathbf{x} - \mathbf{b}^T \mathbf{y}$$

Classic residual associated to the KKT system, for $\alpha > 0$

$$r^p_{\text{KKT}}(\mathbf{x}, \mathbf{y}, \boldsymbol{\pi}) = \|\alpha r_{\text{stat}} + r_{\text{comp}}\|_p$$

$$\text{with} \quad \begin{aligned} r_{\text{stat}}(\mathbf{x}, \mathbf{y}, \boldsymbol{\pi}) &= F(\mathbf{x})^T \mathbf{x} - \mathbf{A}^T \mathbf{y} - \boldsymbol{\pi} \\ r_{\text{comp}}(\mathbf{x}, \boldsymbol{\pi}) &= \mathbf{x} \circ \boldsymbol{\pi} = (x_i \pi_i)_{i=1}^n \end{aligned}$$

Inverse problem with full data

Notation: $\text{MP}(\mathcal{K}, F)$ both refers to $\text{VI}(\mathcal{K}, F)$ and $\text{OP}(\mathcal{K}, f)$

Given $\mathbf{x}^{\text{obs}}$ (nearly) optimal for $\text{MP}(\mathcal{K}, F)$, the inverse problem is convex:[1]

$$\begin{aligned} \min_{\mathbf{y},\boldsymbol{\theta}} \quad & r(\mathbf{x}^{\text{obs}}, \mathbf{y}, \boldsymbol{\theta}) \\ \text{s.t.} \quad & \text{dual feasibility} \\ & \boldsymbol{\theta} \in \Theta \end{aligned}$$

[1] Bertsimas et al. (2014) and Keshavarz, Wang, and Boyd (2011)

Formulation of the inverse problem with missing data

Notation: $\text{MP}(\mathcal{K}, F)$ both refers to $\text{VI}(\mathcal{K}, F)$ and $\text{OP}(\mathcal{K}, f)$

$$\begin{aligned} \min_{\mathbf{y},\boldsymbol{\theta}} \quad & r(\mathbf{x}^{\text{obs}}, \mathbf{y}, \boldsymbol{\theta}) \\ \text{s.t.} \quad & \text{dual feasibility} \\ & \boldsymbol{\theta} \in \Theta \end{aligned}$$

Formulation of the inverse problem with missing data

Notation: $\text{MP}(\mathcal{K}, F)$ both refers to $\text{VI}(\mathcal{K}, F)$ and $\text{OP}(\mathcal{K}, f)$

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} \quad & r(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) \\ \text{s.t.} \quad & \text{dual feasibility} \\ & \mathbf{x} = \mathbf{x}^{\text{obs}} \\ & \boldsymbol{\theta} \in \Theta \end{aligned}$$

Formulation of the inverse problem with missing data

Notation: $\mathrm{MP}(\mathcal{K}, F)$ both refers to $\mathrm{VI}(\mathcal{K}, F)$ and $\mathrm{OP}(\mathcal{K}, f)$

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} \quad & r(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) \\ \text{s.t.} \quad & \text{dual feasibility} \\ & \mathbf{H}\mathbf{x} = \mathbf{z}^{\text{obs}} \\ & \boldsymbol{\theta} \in \Theta \end{aligned}$$

Formulation of the inverse problem with missing data

Notation: $\mathrm{MP}(\mathcal{K}, F)$ both refers to $\mathrm{VI}(\mathcal{K}, F)$ and $\mathrm{OP}(\mathcal{K}, f)$

$$\begin{array}{ll} \min\limits_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} & r(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) \\ \text{s.t.} & \text{dual feasibility} \\ & \mathbf{Hx} = \mathbf{z}^{\text{obs}} \\ & \boldsymbol{\theta} \in \Theta \end{array}$$

- impose **primal feasibility** on the induced response
- formulation **robust to outliers** in the observations
- and $\mathbf{Ax} = \mathbf{b}$, $\mathbf{Hx} = \mathbf{z}$ might be infeasible because of noise

Formulation of the inverse problem with missing data

Notation: $\mathrm{MP}(\mathcal{K}, F)$ both refers to $\mathrm{VI}(\mathcal{K}, F)$ and $\mathrm{OP}(\mathcal{K}, f)$

$$\begin{array}{ll} \min\limits_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} & r(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) \\ \text{s.t.} & \text{dual feasibility} \\ & \mathbf{H}\mathbf{x} = \mathbf{z}^{\text{obs}} \\ & \boldsymbol{\theta} \in \Theta \end{array}$$

- impose **primal feasibility** on the induced response
- formulation **robust to outliers** in the observations
- and $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{H}\mathbf{x} = \mathbf{z}$ might be infeasible because of noise

$$\begin{array}{ll} \min\limits_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} & [r(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}),\ \phi(\mathbf{H}\mathbf{x} - \mathbf{z}^{\text{obs}})]^T \\ \text{s.t.} & \text{primal feasibility} \\ & \text{dual feasibility} \\ & \boldsymbol{\theta} \in \Theta \end{array}$$

Remark: replacing ϕ by general objective g gives a novel single-level formulation of bilevel programs:

$$\min_{\mathbf{x},\boldsymbol{\theta}\in\Theta} g(\mathbf{x}, \boldsymbol{\theta}) \quad \text{s.t.} \quad \mathbf{x} \text{ is solution to } \mathrm{MP}(\mathcal{K}, F(\cdot, \boldsymbol{\theta}))$$

Outline

Inverse problem with missing data

Formulation as a Pareto optimization problem

Theoretical results and implementation

Bounds on residuals

Theorem (Bertsimas et al. 2014)

Suppose primal feasibility and dual feasibility hold. Then
$r_{\mathrm{PD}} \leq \epsilon \quad \Longleftrightarrow \quad r_{\mathrm{VI}} \leq \epsilon \quad \Longrightarrow \quad r_{\mathrm{OP}} \leq \epsilon$

Theorem (Thai and Bayen 2014)

Suppose primal and dual feasibilities hold. Then $\forall\, p \geq 1,\ \alpha > 0$
$r_{\mathrm{VI}} \leq \epsilon \Longrightarrow r_{\mathrm{KKT}}^{p} \leq \epsilon$.
Reciprocally, $r_{\mathrm{KKT}}^{p} \leq \epsilon \Longrightarrow r_{\mathrm{VI}} = O(\epsilon \|\mathbf{x}\| n^{1-\frac{1}{p}})$

- Tight bounds.
- r_{VI} and r_{KKT} define different metrics.

Bounds on residuals for strongly monotone functions

Definition: strong monotonicity

A map $F : \mathbb{R}^n \to \mathbb{R}^n$ is strongly monotone if $\exists\, m > 0$ such that $(F(\mathbf{x}) - F(\mathbf{y}))^T(\mathbf{x} - \mathbf{y}) \geq m\|\mathbf{x} - \mathbf{y}\|_2^2, \quad \forall\, \mathbf{x},\, \mathbf{y} \in \mathcal{K}$

- equivalent to strong convexity of f when $\nabla f = F$.
- unique solution $\mathbf{x}^\star$ to $\mathrm{VI}(\mathcal{K}, F)$, resp. $\mathrm{OP}(\mathcal{K}, f)$.

Bounds on residuals for strongly monotone functions

Definition: strong monotonicity

A map $F : \mathbb{R}^n \to \mathbb{R}^n$ is strongly monotone if $\exists\, m > 0$ such that $(F(\mathbf{x}) - F(\mathbf{y}))^T(\mathbf{x} - \mathbf{y}) \geq m\|\mathbf{x} - \mathbf{y}\|_2^2, \quad \forall\, \mathbf{x}, \mathbf{y} \in \mathcal{K}$

- equivalent to strong convexity of f when $\nabla f = F$.
- unique solution $\mathbf{x}^\star$ to $\mathrm{VI}(\mathcal{K}, F)$, resp. $\mathrm{OP}(\mathcal{K}, f)$.

Theorem (Pang 1996)

Suppose F strongly monotone, primal and dual feasibilities, then $r_{\mathrm{PD}} \leq \epsilon \implies \|\mathbf{x} - \mathbf{x}^\star\|_2 \leq \sqrt{\epsilon/m}$

Bounds on residuals for strongly monotone functions

Definition: strong monotonicity

A map $F : \mathbb{R}^n \to \mathbb{R}^n$ is strongly monotone if $\exists\, m > 0$ such that $(F(\mathbf{x}) - F(\mathbf{y}))^T(\mathbf{x} - \mathbf{y}) \geq m\|\mathbf{x} - \mathbf{y}\|_2^2, \quad \forall\, \mathbf{x}, \mathbf{y} \in \mathcal{K}$

- equivalent to strong convexity of f when $\nabla f = F$.
- unique solution $\mathbf{x}^\star$ to $\mathrm{VI}(\mathcal{K}, F)$, resp. $\mathrm{OP}(\mathcal{K}, f)$.

Theorem (Pang 1996)

Suppose F strongly monotone, primal and dual feasibilities, then $r_{\mathrm{PD}} \leq \epsilon \implies \|\mathbf{x} - \mathbf{x}^\star\|_2 \leq \sqrt{\epsilon/m}$

Theorem (Thai and Bayen 2014)

Suppose F strongly monotone, primal and dual feasibilities, then

$$r_{\mathrm{KKT}} \leq \epsilon \implies \|\mathbf{x} - \mathbf{x}^\star\|_2 \leq O\left(\sqrt{\epsilon\|\mathbf{x}\|_\infty n^{1-\frac{1}{p}}/m}\right)$$

Finding the optimal Pareto point in one shot

$$
\begin{aligned}
\min_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} \quad & [r(\mathbf{x},\mathbf{y},\boldsymbol{\theta}),\ \phi(\mathbf{H}\mathbf{x} - \mathbf{z}^{\text{obs}})]^T \\
\text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \\
& \mathbf{A}^T\mathbf{y} \preceq F(\mathbf{x}|\boldsymbol{\theta}) \\
& \boldsymbol{\theta} \in \Theta
\end{aligned}
$$

Finding the optimal Pareto point in one shot

$$
\begin{aligned}
\min_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} \quad & w_{\text{mp}}\, r(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) + w_{\text{obs}}\, \phi(\mathbf{H}\mathbf{x} - \mathbf{z}^{\text{obs}}) \\
\text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \\
& \mathbf{A}^T \mathbf{y} \preceq F(\mathbf{x}|\boldsymbol{\theta}) \\
& \boldsymbol{\theta} \in \Theta
\end{aligned}
$$

Finding the optimal Pareto point in one shot

$$\begin{aligned}\min_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} \quad & w_{\text{mp}}\, r(\mathbf{x},\mathbf{y},\boldsymbol{\theta}) + w_{\text{obs}}\, \phi(\mathbf{H}\mathbf{x} - \mathbf{z}^{\text{obs}}) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \\ & \mathbf{A}^T\mathbf{y} \preceq F(\mathbf{x}|\boldsymbol{\theta}) \\ & \boldsymbol{\theta} \in \Theta\end{aligned}$$

Classic methodology to explore the Pareto curve

1. Normalize: $\tilde{r} := r/r^{\text{max}}$ and $\tilde{\phi} = \phi/\phi^{\text{max}}$.
2. Solve with $w_{\text{mp}} + w_{obs} = 1$, $w_{\text{mp}} \in \{10^{-2}, 10^{-1}, 0.5, 0.9, 0.99\}$.
3. Check values of the residuals r and ϕ.

Finding the optimal Pareto point in one shot

$$\begin{array}{ll} \min_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} & w_{\text{mp}}\, r(\mathbf{x},\mathbf{y},\boldsymbol{\theta}) + w_{\text{obs}}\, \phi(\mathbf{H}\mathbf{x} - \mathbf{z}^{\text{obs}}) \\ \text{s.t.} & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \\ & \mathbf{A}^T\mathbf{y} \preceq F(\mathbf{x}|\boldsymbol{\theta}) \\ & \boldsymbol{\theta} \in \Theta \end{array}$$

Classic methodology to explore the Pareto curve

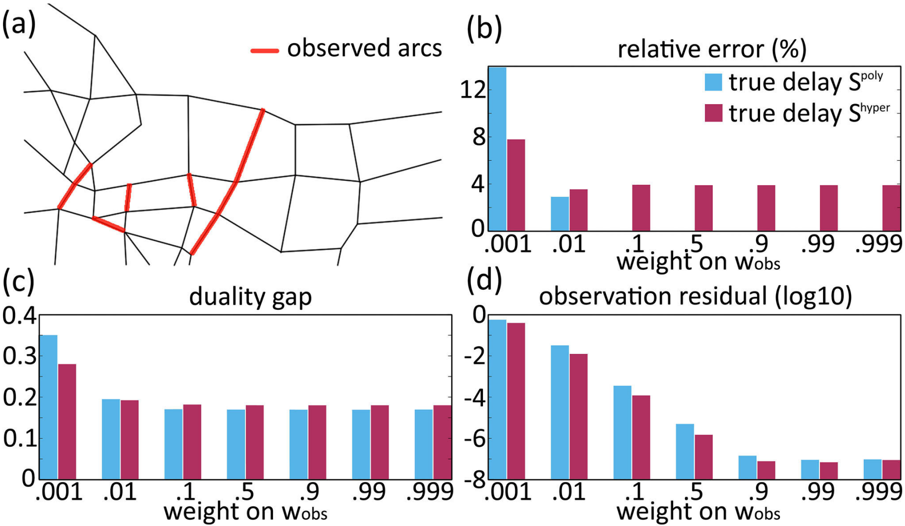
1. Normalize: $\tilde{r} := r/r^{\max}$ and $\tilde{\phi} = \phi/\phi^{\max}$.
2. Solve with $w_{\text{mp}} + w_{obs} = 1$, $w_{\text{mp}} \in \{10^{-2}, 10^{-1}, 0.5, 0.9, 0.99\}$.
3. Check values of the residuals r and ϕ.

With noiseless data, sufficient to solve one program with $w_{\text{obs}} \approx 1$

Theorem 9 (Thai and Bayen 2014)

If $\exists\, \mathbf{x},\ \mathbf{y},\ \boldsymbol{\theta}$ such that $\quad r(\mathbf{x},\mathbf{y},\boldsymbol{\theta}) \leq \epsilon, \quad \mathbf{H}\mathbf{x} = \mathbf{z}^{\text{obs}}$
Then a solution $(\mathbf{x}^\star, \mathbf{y}^\star, \boldsymbol{\theta}^\star)$ to the weighted sum method is such that
$r(\mathbf{x}^\star, \mathbf{y}^\star, \boldsymbol{\theta}^\star) \leq \epsilon, \quad \phi(\mathbf{H}\mathbf{x}^\star - \mathbf{z}^{\text{obs}}) \to 0 \quad \text{as} \quad w_{\text{obs}} \to 1$

Numerical experiments: weighted sum method



Parallelization over multiple observations

- Given pairs $(\mathbf{z}_j^{\text{obs}}, \mathcal{K}_j)$ for $j = 1, \cdots, N$
- $\mathcal{K}_j = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}_j,\ \mathbf{x} \succeq 0\}$ encodes a specific configuration
- $\mathbf{x}_j$ is the resulting optimal response, but only observe $\mathbf{z}_j^{\text{obs}} = \mathbf{H}\mathbf{x}_j$
- Find $\boldsymbol{\theta}$ and $\{\mathbf{x}_j\}_j$ solution to $\text{VI}(\mathcal{K}_j, F(\cdot|\boldsymbol{\theta}))$, $\mathbf{H}\mathbf{x}_j = \mathbf{z}_j^{\text{obs}}\ \forall\, j$

Parallelization over multiple observations

- Given pairs $(\mathbf{z}_j^{\text{obs}}, \mathcal{K}_j)$ for $j = 1, \cdots, N$
- $\mathcal{K}_j = \{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b}_j,\ \mathbf{x} \succeq 0\}$ encodes a specific configuration
- $\mathbf{x}_j$ is the resulting optimal response, but only observe $\mathbf{z}_j^{\text{obs}} = \mathbf{H}\mathbf{x}_j$
- Find $\boldsymbol{\theta}$ and $\{\mathbf{x}_j\}_j$ solution to $\text{VI}(\mathcal{K}_j, F(\cdot|\boldsymbol{\theta}))$, $\mathbf{H}\mathbf{x}_j = \mathbf{z}_j^{\text{obs}}\ \forall j$

$$\begin{aligned}
\min_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} \quad & w_{\text{mp}}\, r(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) + w_{\text{obs}}\, \phi(\mathbf{H}\mathbf{x} - \mathbf{z}^{\text{obs}}) \\
\text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0 \\
& \mathbf{A}^T\mathbf{y} \preceq F(\mathbf{x}|\boldsymbol{\theta}) \\
& \boldsymbol{\theta} \in \Theta
\end{aligned}$$

Parallelization over multiple observations

- Given pairs $(\mathbf{z}_j^{\text{obs}}, \mathcal{K}_j)$ for $j = 1, \cdots, N$
- $\mathcal{K}_j = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}_j,\ \mathbf{x} \succeq 0\}$ encodes a specific configuration
- $\mathbf{x}_j$ is the resulting optimal response, but only observe $\mathbf{z}_j^{\text{obs}} = \mathbf{H}\mathbf{x}_j$
- Find $\boldsymbol{\theta}$ and $\{\mathbf{x}_j\}_j$ solution to $\text{VI}(\mathcal{K}_j, F(\cdot|\boldsymbol{\theta}))$, $\mathbf{H}\mathbf{x}_j = \mathbf{z}_j^{\text{obs}}\ \forall j$

$$
\begin{aligned}
\min_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} \quad & w_{\text{mp}} \textstyle\sum_j r(\mathbf{x}_j, \mathbf{y}_j, \boldsymbol{\theta}) + w_{\text{obs}} \textstyle\sum_j \phi(\mathbf{H}\mathbf{x}_j - \mathbf{z}_j{}^{\text{obs}}) \\
\text{s.t.} \quad & \mathbf{A}_j\mathbf{x}_j = \mathbf{b}_j,\ \mathbf{x}_j \succeq 0, \quad j = 1, \cdots, N \\
& \mathbf{A}_j{}^T\mathbf{y}_j \preceq F(\mathbf{x}_j|\boldsymbol{\theta}) \quad j = 1, \cdots, N \\
& \boldsymbol{\theta} \in \Theta
\end{aligned}
$$

Parallelization over multiple observations

- Given pairs $(\mathbf{z}_j^{\text{obs}}, \mathcal{K}_j)$ for $j = 1, \cdots, N$
- $\mathcal{K}_j = \{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b}_j,\ \mathbf{x} \succeq 0\}$ encodes a specific configuration
- $\mathbf{x}_j$ is the resulting optimal response, but only observe $\mathbf{z}_j^{\text{obs}} = \mathbf{H}\mathbf{x}_j$
- Find $\boldsymbol{\theta}$ and $\{\mathbf{x}_j\}_j$ solution to $\text{VI}(\mathcal{K}_j, F(\cdot|\boldsymbol{\theta}))$, $\mathbf{H}\mathbf{x}_j = \mathbf{z}_j^{\text{obs}}\ \forall j$

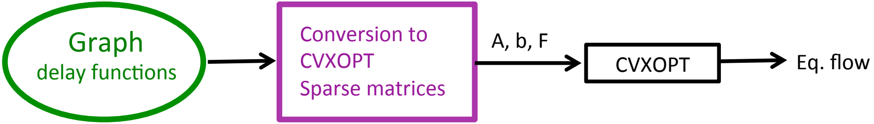
$$\begin{array}{ll}\min\limits_{\mathbf{x},\mathbf{y},\boldsymbol{\theta}} & w_{\text{mp}} \sum_j r(\mathbf{x}_j, \mathbf{y}_j, \boldsymbol{\theta}) + w_{\text{obs}} \sum_j \phi(\mathbf{H}\mathbf{x}_j - \mathbf{z}_j{}^{\text{obs}}) \\ \text{s.t.} & \mathbf{A}_j\mathbf{x}_j = \mathbf{b}_j,\ \mathbf{x}_j \succeq 0, \quad j = 1, \cdots, N \\ & \mathbf{A}_j{}^T\mathbf{y}_j \preceq F(\mathbf{x}_j|\boldsymbol{\theta}) \quad j = 1, \cdots, N \\ & \boldsymbol{\theta} \in \Theta \end{array}$$

$\boldsymbol{\theta}$ is the common structural parameter. When fixed, **parallelizable**:

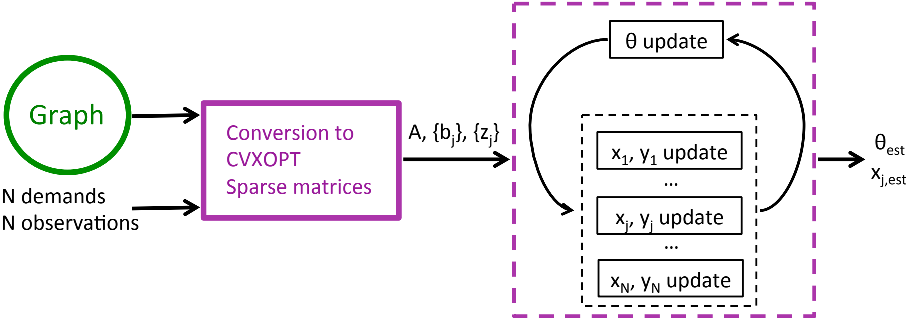
$$\begin{array}{ll}\min\limits_{\mathbf{x}_j,\mathbf{y}_j} & w_{\text{mp}}\, r(\mathbf{x}_j, \mathbf{y}_j, \boldsymbol{\theta}) + w_{\text{obs}}\, \phi(\mathbf{H}\mathbf{x}_j - \mathbf{z}_j^{\text{obs}}) \\ \text{s.t.} & \mathbf{A}_j\mathbf{x}_j = \mathbf{b}_j,\ \mathbf{x}_j \succeq 0 \\ & \mathbf{A}_j^T\mathbf{y}_j \preceq F(\mathbf{x}_j|\boldsymbol{\theta}) \end{array} \qquad \text{for } j = 1, \cdots, N$$

Algorithm: update cyclically convex blocks $\{\mathbf{x}_j\}_{j=1}^N$, $\{\mathbf{y}_j\}_{j=1}^N$, $\boldsymbol{\theta}$

Implementation of the forward and reverse solvers



Forward solver



Reverse solver

Publications

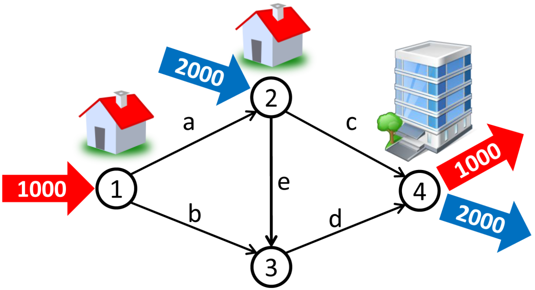
- J. Thai, R. Hariss, A. Bayen, Approximate Bilevel Programming via Pareto Optimization for Imputation and Control of Optimization and Equilibrium models, *accepted, ECC2015*
- J. Thai, R. Hariss, A. Bayen, A Multi-Convex approach to Latency Inference and Control in Traffic Equilibria from Sparse data, *accepted, ACC2015*

Future works

- Data driven re-estimation of the BPR function
- Estimation robust to attacks using the ℓ_1 norm
- Model fitting to mimic more complex behaviors and for efficient re-routing
- Large-scale implementation of the inverse problem with GPS and cellular data

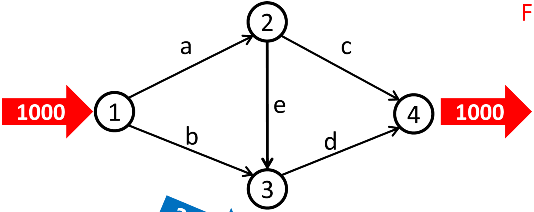
Appendix: traffic assignment problem

Morning commute example for the traffic assignment problem



- $\mathcal{A}$ = arc set = $\{a, b, c, d, e\}$
- $\mathcal{N}$ = node set = $\{1, 2, 3, 4\}$
- Commodity 1: $c_1 = (1 \rightarrow 4, 1000)$ "routing 1000 veh/h from 1 to 4"
- Commodity 2: $c_2 = (2 \rightarrow 4, 2000)$ "routing 2000 veh/h from 2 to 4"
- $\mathcal{C}$ = commodity set = $\{c_1, c_2\}$

Morning commute example for the traffic assignment problem

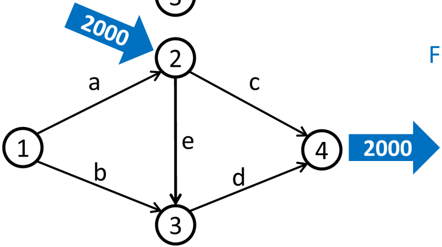


Flow conservation for commodity 1

$$\begin{aligned} -x_a - x_b &= -1000 \\ x_a - x_c - x_e &= 0 \\ x_b + x_e - x_d &= 0 \\ x_c + x_d &= 1000 \end{aligned}$$

nonnegative flows

$\Rightarrow$ $\mathbf{N\, x_1 = b_1}$, $\mathbf{x_1 \geq 0}$



Flow conservation for commodity 2

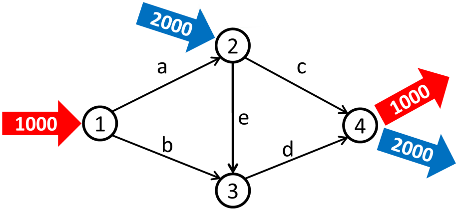
$$\begin{aligned} -x_a - x_b &= 0 \\ x_a - x_c - x_e &= -2000 \\ x_b + x_e - x_d &= 0 \\ x_c + x_d &= 2000 \end{aligned}$$

nonnegative flows

$\Rightarrow$ $\mathbf{N\, x_2 = b_2}$, $\mathbf{x_2 \geq 0}$

- $\mathcal{A} = \{a, b, c, d, e\}$, $|\mathcal{A}| = 5$, $\mathcal{C} = \{1 \to 4,\ 2 \to 4\}$, $|\mathcal{C}| = 2$
- commodity flow vectors: $\mathbf{x}_1 \in \mathbb{R}^{|\mathcal{A}|}$, $\mathbf{x}_2 \in \mathbb{R}^{|\mathcal{A}|}$
- overall flow vector: $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{R}^{|\mathcal{A}| \times |\mathcal{C}|}$
- aggregate flow vector: $\mathbf{v} = \mathbf{x}_1 + \mathbf{x}_2 \in \mathbb{R}^{|\mathcal{A}|}$

Morning commute example for the traffic assignment problem



- Feasible set:

$$\mathcal{K} = \{\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) \mid \mathbf{N}\mathbf{x}_1 = \mathbf{b}_1, \mathbf{x}_1 \succeq 0, \mathbf{N}\mathbf{x}_2 = \mathbf{b}_2, \mathbf{x}_2 \succeq 0\}$$

- Delay map $S : \mathbb{R}^{|\mathcal{A}|} \to \mathbb{R}^{|\mathcal{A}|}$ w.r.t. aggregate flow $\mathbf{v} = \mathbf{x}_1 + \mathbf{x}_2$

$$S(\mathbf{v}) = (\, s_a(v_a), s_b(v_b), s_c(v_c), s_d(v_d), s_e(v_e)\,) \in \mathbb{R}^{|\mathcal{A}|}$$

- $\mathbf{x}^\star \in \mathcal{K}$ is a Nash eq. if $\forall\, \mathbf{x} \in \mathcal{K}$, the associated aggregate flows $\mathbf{v}^\star$, $\mathbf{v}$ are such that

$$v_a^\star\, s_a(v_a^\star) + \cdots + v_e^\star\, s_e(v_e^\star) \leq v_a\, s_a(v_a^\star) + \cdots + v_e\, s_e(v_e^\star)$$

General traffic assignment problem

- commodity flow vectors: $\mathbf{x}_k \in \mathbb{R}^{|\mathcal{A}|}$ for all $k \in \mathcal{C}$

General traffic assignment problem

- commodity flow vectors: $\mathbf{x}_k \in \mathbb{R}^{|\mathcal{A}|}$ for all $k \in \mathcal{C}$
- overall flow vector: $\mathbf{x} = (\mathbf{x}_k)_{k \in \mathcal{C}} \in \mathbb{R}^{|\mathcal{A}| \times |\mathcal{C}|}$

General traffic assignment problem

- commodity flow vectors: $\mathbf{x}_k \in \mathbb{R}^{|\mathcal{A}|}$ for all $k \in \mathcal{C}$
- overall flow vector: $\mathbf{x} = (\mathbf{x}_k)_{k \in \mathcal{C}} \in \mathbb{R}^{|\mathcal{A}| \times |\mathcal{C}|}$
- aggregate flow vector: $\mathbf{v} = \sum_{k \in \mathcal{C}} \mathbf{x}_k = \mathbf{Z}\mathbf{x} \in \mathbb{R}^{|\mathcal{A}|}$

General traffic assignment problem

- commodity flow vectors: $\mathbf{x}_k \in \mathbb{R}^{|\mathcal{A}|}$ for all $k \in \mathcal{C}$
- overall flow vector: $\mathbf{x} = (\mathbf{x}_k)_{k \in \mathcal{C}} \in \mathbb{R}^{|\mathcal{A}| \times |\mathcal{C}|}$
- aggregate flow vector: $\mathbf{v} = \sum_{k \in \mathcal{C}} \mathbf{x}_k = \mathbf{Z}\mathbf{x} \in \mathbb{R}^{|\mathcal{A}|}$
- feasible set:
$$\begin{aligned} \mathcal{K} &= \{\mathbf{x} = (\mathbf{x}_k)_{k \in \mathcal{C}} \mid \mathbf{N}\mathbf{x}_k = \mathbf{b}_k,\ \mathbf{x}_k \succeq 0,\ \forall\, k \in \mathcal{C}\} \\ &= \{\mathbf{x} \in \mathbb{R}^{|\mathcal{A}| \times |\mathcal{C}|} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0\} \end{aligned}$$

General traffic assignment problem

- commodity flow vectors: $\mathbf{x}_k \in \mathbb{R}^{|\mathcal{A}|}$ for all $k \in \mathcal{C}$
- overall flow vector: $\mathbf{x} = (\mathbf{x}_k)_{k\in\mathcal{C}} \in \mathbb{R}^{|\mathcal{A}|\times|\mathcal{C}|}$
- aggregate flow vector: $\mathbf{v} = \sum_{k\in\mathcal{C}} \mathbf{x}_k = \mathbf{Z}\mathbf{x} \in \mathbb{R}^{|\mathcal{A}|}$
- feasible set: $\mathcal{K} = \{\mathbf{x} = (\mathbf{x}_k)_{k\in\mathcal{C}} \,|\, \mathbf{N}\mathbf{x}_k = \mathbf{b}_k,\ \mathbf{x}_k \succeq 0,\ \forall\, k \in \mathcal{C}\}$
$$= \{\mathbf{x} \in \mathbb{R}^{|\mathcal{A}|\times|\mathcal{C}|} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0\}$$
- Delay map $S : \mathbb{R}^{|\mathcal{A}|} \to \mathbb{R}^{|\mathcal{A}|}$ such that $S(\mathbf{v}) = (s_a(v_a))_{a\in\mathcal{A}}$

General traffic assignment problem

- commodity flow vectors: $\mathbf{x}_k \in \mathbb{R}^{|\mathcal{A}|}$ for all $k \in \mathcal{C}$
- overall flow vector: $\mathbf{x} = (\mathbf{x}_k)_{k\in\mathcal{C}} \in \mathbb{R}^{|\mathcal{A}|\times|\mathcal{C}|}$
- aggregate flow vector: $\mathbf{v} = \sum_{k\in\mathcal{C}} \mathbf{x}_k = \mathbf{Z}\mathbf{x} \in \mathbb{R}^{|\mathcal{A}|}$
- feasible set: $\mathcal{K} = \{\mathbf{x} = (\mathbf{x}_k)_{k\in\mathcal{C}} \,|\, \mathbf{N}\mathbf{x}_k = \mathbf{b}_k,\, \mathbf{x}_k \succeq 0,\, \forall\, k \in \mathcal{C}\}$
  $= \{\mathbf{x} \in \mathbb{R}^{|\mathcal{A}|\times|\mathcal{C}|} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b},\, \mathbf{x} \succeq 0\}$
- Delay map $S : \mathbb{R}^{|\mathcal{A}|} \to \mathbb{R}^{|\mathcal{A}|}$ such that $S(\mathbf{v}) = (s_a(v_a))_{a\in\mathcal{A}}$
- Nash equilibrium: $\mathbf{x}^\star \in \mathcal{K}$ such that for all $\mathbf{x} \in \mathcal{K}$

$$\begin{aligned}
\sum_{a\in\mathcal{A}} v_a^\star\, s_a(v_a^\star) \le \sum_{a\in\mathcal{A}} v_a\, s_a(v_a^\star) \quad &\iff \quad S(\mathbf{v}^\star)^T\mathbf{v}^\star \le S(\mathbf{v}^\star)^T\mathbf{v} \\
&\iff \quad S(\mathbf{Z}\mathbf{x}^\star)^T\mathbf{Z}\mathbf{x}^\star \le S(\mathbf{Z}\mathbf{x}^\star)^T\mathbf{Z}\mathbf{x} \\
&\iff \quad F(\mathbf{x}^\star)^T\mathbf{x}^\star \le F(\mathbf{x}^\star)^T\mathbf{x}
\end{aligned}$$

$\implies$ Nash eq. $=$ solution to a VI with $F(\mathbf{x}) = \mathbf{Z}^T S(\mathbf{Z}\mathbf{x})$

General traffic assignment problem

- commodity flow vectors: $\mathbf{x}_k \in \mathbb{R}^{|\mathcal{A}|}$ for all $k \in \mathcal{C}$
- overall flow vector: $\mathbf{x} = (\mathbf{x}_k)_{k\in\mathcal{C}} \in \mathbb{R}^{|\mathcal{A}|\times|\mathcal{C}|}$
- aggregate flow vector: $\mathbf{v} = \sum_{k\in\mathcal{C}} \mathbf{x}_k = \mathbf{Z}\mathbf{x} \in \mathbb{R}^{|\mathcal{A}|}$
- feasible set: $\mathcal{K} = \{\mathbf{x} = (\mathbf{x}_k)_{k\in\mathcal{C}} \mid \mathbf{N}\mathbf{x}_k = \mathbf{b}_k,\, \mathbf{x}_k \succeq 0,\, \forall\, k \in \mathcal{C}\}$
  $= \{\mathbf{x} \in \mathbb{R}^{|\mathcal{A}|\times|\mathcal{C}|} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\, \mathbf{x} \succeq 0\}$
- Delay map $S : \mathbb{R}^{|\mathcal{A}|} \to \mathbb{R}^{|\mathcal{A}|}$ such that $S(\mathbf{v}) = (s_a(v_a))_{a\in\mathcal{A}}$
- Nash equilibrium: $\mathbf{x}^\star \in \mathcal{K}$ such that for all $\mathbf{x} \in \mathcal{K}$

$$\sum_{a\in\mathcal{A}} v_a^\star\, s_a(v_a^\star) \le \sum_{a\in\mathcal{A}} v_a\, s_a(v_a^\star) \iff S(\mathbf{v}^\star)^T\mathbf{v}^\star \le S(\mathbf{v}^\star)^T\mathbf{v}$$
$$\iff S(\mathbf{Z}\mathbf{x}^\star)^T\mathbf{Z}\mathbf{x}^\star \le S(\mathbf{Z}\mathbf{x}^\star)^T\mathbf{Z}\mathbf{x}$$
$$\iff F(\mathbf{x}^\star)^T\mathbf{x}^\star \le F(\mathbf{x}^\star)^T\mathbf{x}$$

$\implies$ Nash eq. $=$ solution to a VI with $F(\mathbf{x}) = \mathbf{Z}^T S(\mathbf{Z}\mathbf{x})$

Definition: variational inequality (VI)

VI$(\mathcal{K}, F)$: find $\mathbf{x}^\star \in \mathcal{K}$ such that $F(\mathbf{x}^\star)^T(\mathbf{x} - \mathbf{x}^\star) \ge 0,\ \forall\, \mathbf{x} \in \mathcal{K}$.

Optimization process and Variational inequality

Theorem 1 (Beckmann et al. 1956)

Suppose the arc delay functions are nonnegative, continuous, monotone, separable. Then the Nash equilibrium is solution of a convex optimization program, denoted $\mathrm{OP}(\mathcal{K}, f)$

$$\min\ f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0$$

Remarks

- The potential f encodes the interaction between players.
- $\mathcal{K} := \{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \succeq 0\}$ encodes the flow conservation.

Optimization process and Variational inequality

Theorem 1 (Beckmann et al. 1956)

Suppose the arc delay functions are nonnegative, continuous, monotone, separable. Then the Nash equilibrium is solution of a convex optimization program, denoted $\mathrm{OP}(\mathcal{K}, f)$

$$\min \; f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{Ax} = \mathbf{b},\; \mathbf{x} \succeq 0$$

Remarks

- The potential f encodes the interaction between players.
- $\mathcal{K} := \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b},\; \mathbf{x} \succeq 0\}$ encodes the flow conservation.

Theorem 2

With $f \in C^1$, $\mathbf{x}^\star \in \mathcal{K}$ is solution iff $\nabla f(\mathbf{x}^\star)^T(\mathbf{u} - \mathbf{x}^\star) \geq 0,\; \forall\, \mathbf{u} \in \mathcal{K}$.

Result from Beckmann: for the map $F(\mathbf{x}) = \mathbf{Z}^T S(\mathbf{Zx})$, $\exists\, f$ convex such that $F = \nabla f$