ECE 598 Computational Power Systems

Inverse optimization for market analytics

Vladimir Dvorkin

University of Michigan

Last lecture recap

- Look around you and form teams of 2 people (1 min)
- Quickly review your notes or the slide deck (1 min)
- Share your **three** personal highlights with your partner (3 min)
- Get iClicker app ready

Inversing ISO's decision-making model

data y $\longrightarrow$

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & f(x, y) \\
\text{subject to} \quad & g(x, y) \leqslant 0 : \lambda
\end{aligned}
$$

$\longrightarrow$ decisions x, λ

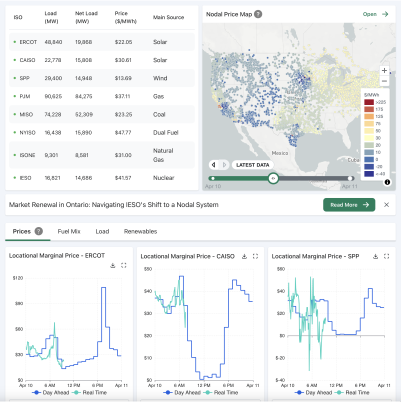
Inversing ISO's decision-making model

data y

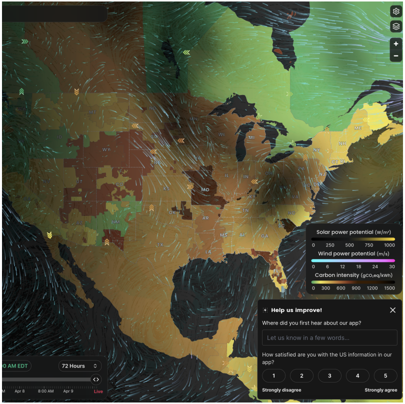
decisions x, λ

- In practice, we do not know optimization underlying the decision-making process
- But we know some data (e.g., weather) and decisions (e.g., dispatch, prices)
- Our goal is to estimate y from decisions x, λ to reconstruct decision-making

Where does the data come from?



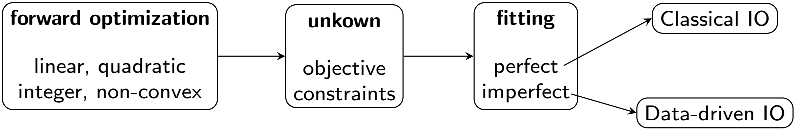
https://www.gridstatus.io/live



app.electricitymaps.com

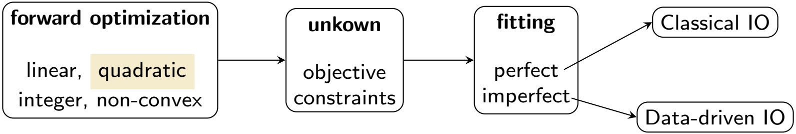
- Abundance of data available in real-time thanks to grid analytics start ups
- Aggregated generation statistics, LMPs, emissions, meteorological data, etc.
- We can rely on these data points to restore the underlying ISO's decision-making

Taxonomy of inverse optimization (IO) problems



- **Forward optimization**: the decision-making model whose parameters we intend to estimate from the optimal decisions
- **Unkown**: parameters of the objective function, constraint function, or both
- **Fitting**: Depending on available data, we may achieve the perfect fit or, at least, maximize a suitable measure of fitness

Taxonomy of inverse optimization (IO) problems



- **Forward optimization**: the decision-making model whose parameters we intend to estimate from the optimal decisions
- **Unkown**: parameters of the objective function, constraint function, or both
- **Fitting**: Depending on available data, we may achieve the perfect fit or, at least, maximize a suitable measure of fitness

Forward optimization (FO) problem

Original optimization problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \mathbf{C}\mathbf{x} + \mathbf{c}^\top \mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \geqslant \mathbf{b} \end{aligned}$$

Forward optimization problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \boldsymbol{\Theta}\mathbf{x} + \boldsymbol{\theta}^\top \mathbf{x} \\ \text{subject to} \quad & \boldsymbol{\Psi}\mathbf{x} \geqslant \boldsymbol{\psi} \end{aligned}$$

- FO replicates the structure of the original problem
- $\boldsymbol{\Theta}$ and $\boldsymbol{\theta}$ are unknown parameters of the objective function
- $\boldsymbol{\Psi}$ and $\boldsymbol{\psi}$ are unknown parameters of the constraint function

Forward optimization (FO) problem

Original optimization problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \mathbf{C}\mathbf{x} + \mathbf{c}^\top \mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \geqslant \mathbf{b} \end{aligned}$$

Forward optimization problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \boldsymbol{\Theta}\mathbf{x} + \boldsymbol{\theta}^\top \mathbf{x} \\ \text{subject to} \quad & \boldsymbol{\Psi}\mathbf{x} \geqslant \boldsymbol{\psi} \end{aligned}$$

- FO replicates the structure of the original problem
- $\boldsymbol{\Theta}$ and $\boldsymbol{\theta}$ are unknown parameters of the objective function
- $\boldsymbol{\Psi}$ and $\boldsymbol{\psi}$ are unknown parameters of the constraint function

- **Estimating objective function**: set $\boldsymbol{\Psi} = \mathbf{A}$ and $\boldsymbol{\psi} = \mathbf{b}$, estimate $\boldsymbol{\Theta}$ and $\boldsymbol{\theta}$
- **Estimating constraint function**: set $\boldsymbol{\Theta} = \mathbf{C}$ and $\boldsymbol{\theta} = \mathbf{c}$, estimate $\boldsymbol{\Psi}$ and $\boldsymbol{\psi}$

Classical IO using bilevel programming

- Second-order cost coefficients in $\mathbf{C}$ and constraint matrix $\mathbf{A}$ are known
- First-order cost coefficients in $\mathbf{c}$ and the right-hand side parameter $\mathbf{b}$ are unknown
- Known optimal primal and dual decisions $\mathbf{x}^\star$ and $\boldsymbol{\lambda}^\star$ from the original problem

Classical IO using bilevel programming

- Second-order cost coefficients in **C** and constraint matrix **A** are known
- First-order cost coefficients in **c** and the right-hand side parameter **b** are unknown
- Known optimal primal and dual decisions $\mathbf{x}^\star$ and $\boldsymbol{\lambda}^\star$ from the original problem

$$\begin{aligned}
\underset{\boldsymbol{\theta},\boldsymbol{\psi}}{\text{minimize}} \quad & \underbrace{\|\mathbf{x}(\boldsymbol{\theta},\boldsymbol{\psi}) - \mathbf{x}^\star\|_p}_{\text{primal error}} + \underbrace{\|\boldsymbol{\lambda}(\boldsymbol{\theta},\boldsymbol{\psi}) - \boldsymbol{\lambda}^\star\|_p}_{\text{dual error}} && \textit{minimize decision error} \\
\text{subject to} \quad & \underline{\boldsymbol{\theta}} \leqslant \boldsymbol{\theta} \leqslant \overline{\boldsymbol{\theta}} && \textit{constrain unknowns} \\
& \underline{\boldsymbol{\psi}} \leqslant \boldsymbol{\psi} \leqslant \overline{\boldsymbol{\psi}} \\
& \mathbf{x}(\boldsymbol{\theta},\boldsymbol{\psi}) \cup \boldsymbol{\lambda}(\boldsymbol{\theta},\boldsymbol{\psi}) \in \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{argmin}} \quad \mathbf{x}^\top \mathbf{C}\mathbf{x} + \boldsymbol{\theta}^\top \mathbf{x} && \textit{feedback from FO} \\
& \qquad\qquad\qquad\qquad \text{subject to} \quad \mathbf{A}\mathbf{x} \geqslant \boldsymbol{\psi} : \boldsymbol{\lambda}
\end{aligned}$$

Classical IO using bilevel programming

- Second-order cost coefficients in **C** and constraint matrix **A** are known
- First-order cost coefficients in **c** and the right-hand side parameter **b** are unknown
- Known optimal primal and dual decisions $\mathbf{x}^\star$ and $\boldsymbol{\lambda}^\star$ from the original problem

$$
\begin{aligned}
\underset{\boldsymbol{\theta},\boldsymbol{\psi}}{\text{minimize}} \quad & \underbrace{\|\mathbf{x}(\boldsymbol{\theta},\boldsymbol{\psi}) - \mathbf{x}^\star\|_p}_{\text{primal error}} + \underbrace{\|\boldsymbol{\lambda}(\boldsymbol{\theta},\boldsymbol{\psi}) - \boldsymbol{\lambda}^\star\|_p}_{\text{dual error}} && \textit{minimize decision error} \\
\text{subject to} \quad & \underline{\boldsymbol{\theta}} \leqslant \boldsymbol{\theta} \leqslant \overline{\boldsymbol{\theta}} && \textit{constrain unknowns} \\
& \underline{\boldsymbol{\psi}} \leqslant \boldsymbol{\psi} \leqslant \overline{\boldsymbol{\psi}} \\
& \mathbf{x}(\boldsymbol{\theta},\boldsymbol{\psi}) \cup \boldsymbol{\lambda}(\boldsymbol{\theta},\boldsymbol{\psi}) \in \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{argmin}} \quad \mathbf{x}^\top \mathbf{C}\mathbf{x} + \boldsymbol{\theta}^\top \mathbf{x} && \textit{feedback from FO} \\
& \qquad\qquad\qquad\qquad\qquad \text{subject to} \quad \mathbf{A}\mathbf{x} \geqslant \boldsymbol{\psi} : \boldsymbol{\lambda}
\end{aligned}
$$

- **Q.** How to solve this problem?

Classical IO using bilevel programming

- Second-order cost coefficients in **C** and constraint matrix **A** are known
- First-order cost coefficients in **c** and the right-hand side parameter **b** are unknown
- Known optimal primal and dual decisions $\mathbf{x}^\star$ and $\boldsymbol{\lambda}^\star$ from the original problem

$$
\begin{aligned}
\underset{\boldsymbol{\theta},\boldsymbol{\psi}}{\text{minimize}} \quad & \underbrace{\|\mathbf{x}(\boldsymbol{\theta},\boldsymbol{\psi}) - \mathbf{x}^\star\|_p}_{\text{primal error}} + \underbrace{\|\boldsymbol{\lambda}(\boldsymbol{\theta},\boldsymbol{\psi}) - \boldsymbol{\lambda}^\star\|_p}_{\text{dual error}} && \textit{minimize decision error} \\
\text{subject to} \quad & \underline{\boldsymbol{\theta}} \leqslant \boldsymbol{\theta} \leqslant \overline{\boldsymbol{\theta}} && \textit{constrain unknowns} \\
& \underline{\boldsymbol{\psi}} \leqslant \boldsymbol{\psi} \leqslant \overline{\boldsymbol{\psi}} \\
& \mathbf{x}(\boldsymbol{\theta},\boldsymbol{\psi}) \cup \boldsymbol{\lambda}(\boldsymbol{\theta},\boldsymbol{\psi}) \in \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{argmin}} \quad \mathbf{x}^\top \mathbf{C}\mathbf{x} + \boldsymbol{\theta}^\top \mathbf{x} && \textit{feedback from FO} \\
& \qquad\qquad\qquad\qquad\qquad \text{subject to} \quad \mathbf{A}\mathbf{x} \geqslant \boldsymbol{\psi} : \boldsymbol{\lambda}
\end{aligned}
$$

- **Q.** How to solve this problem?
- For $p = \{1, \infty\}$, solve the IO as mixed-integer linear program in a single shot
- For $p = 2$, we can also iteratively solve the IO using gradient descent
- See Lecture 10 for solution strategies

Data-driven IO

- In classic IO, we only relied on a single observation $\mathbf{x}^\star$ and $\boldsymbol{\lambda}^\star$
- Data-driven IO acts on the history of *n* observations

$$(\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star), \ldots, (\mathbf{x}_n^\star, \boldsymbol{\lambda}_n^\star)$$

- Used to estimate static parameters (e.g., $\mathbf{c}, \mathbf{b}$) from repeated observations
- or *data generating* models (e.g., forecast models) that produce decision inputs

Data-driven IO for static parameter estimation

- Given a history of n observations $(\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star)$, find the right-hand side parameter $\mathbf{b}$
- Optimized unknown parameter $\boldsymbol{\psi}$ on all n observations

$$
\begin{aligned}
\underset{\boldsymbol{\psi}}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=1}^{n}\left(\|\mathbf{x}_i(\boldsymbol{\psi}) - \mathbf{x}_i^\star\|_p + \|\boldsymbol{\lambda}_i(\boldsymbol{\psi}) - \boldsymbol{\lambda}_i^\star\|_p\right) && \textit{minimize decision error} \\
\text{subject to} \quad & \underline{\boldsymbol{\psi}} \leqslant \boldsymbol{\psi} \leqslant \overline{\boldsymbol{\psi}} \\
& \mathbf{x}_i(\boldsymbol{\psi}) \cup \boldsymbol{\lambda}_i(\boldsymbol{\psi}) \in \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{argmin}} \quad \mathbf{x}^\top \mathbf{C}_i \mathbf{x} + \mathbf{c}_i^\top \mathbf{x} && \textit{FO for every } i \\
& \qquad\qquad\qquad\qquad \text{subject to} \quad \mathbf{A}_i \mathbf{x} \geqslant \boldsymbol{\psi} : \boldsymbol{\lambda}
\end{aligned}
$$

Data-driven IO for static parameter estimation

- Given a history of n observations $(\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star)$, find the right-hand side parameter $\mathbf{b}$
- Optimized unknown parameter $\boldsymbol{\psi}$ on all n observations

$$
\begin{aligned}
\underset{\boldsymbol{\psi}}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=1}^{n}\left(\|\mathbf{x}_i(\boldsymbol{\psi}) - \mathbf{x}_i^\star\|_p + \|\boldsymbol{\lambda}_i(\boldsymbol{\psi}) - \boldsymbol{\lambda}_i^\star\|_p\right) && \textit{minimize decision error} \\
\text{subject to} \quad & \underline{\boldsymbol{\psi}} \leqslant \boldsymbol{\psi} \leqslant \overline{\boldsymbol{\psi}} \\
& \mathbf{x}_i(\boldsymbol{\psi}) \cup \boldsymbol{\lambda}_i(\boldsymbol{\psi}) \in \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{argmin}} \quad \mathbf{x}^\top \mathbf{C}_i \mathbf{x} + \mathbf{c}_i^\top \mathbf{x} && \textit{FO for every } i \\
& \qquad\qquad\qquad\qquad \text{subject to} \quad \mathbf{A}_i \mathbf{x} \geqslant \boldsymbol{\psi} : \boldsymbol{\lambda}
\end{aligned}
$$

- **Q.** How to solve the problem when n is large?

Data-driven IO for data generating models

- Consider that the right-hand side parameter $\mathbf{b}$ is variable
- Suppose it comes from a prediction model:

$$\widehat{\mathbf{b}} = \mathbf{B}\widehat{\boldsymbol{\varphi}}$$

  where $\widehat{\boldsymbol{\varphi}}$ is the feature vector, and $\mathbf{B}$ is the matrix of model weights
- Feature $\widehat{\boldsymbol{\varphi}}$ should correlate variable with $\widehat{\mathbf{b}}$ (e.g., as air temperature for loads)
- Given a history of observations $\{(\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star, \boldsymbol{\varphi}_i)\}_{i=1}^n$, find model weights $\mathbf{B}$

Data-driven IO for data generating models

- Consider that the right-hand side parameter $\mathbf{b}$ is variable
- Suppose it comes from a prediction model:

$$\widehat{\mathbf{b}} = \mathbf{B}\widehat{\varphi}$$

  where $\widehat{\varphi}$ is the feature vector, and $\mathbf{B}$ is the matrix of model weights
- Feature $\widehat{\varphi}$ should correlate variable with $\widehat{\mathbf{b}}$ (e.g., as air temperature for loads)
- Given a history of observations $\{(\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star, \boldsymbol{\varphi}_i)\}_{i=1}^n$, find model weights $\mathbf{B}$

$$\begin{aligned}
\underset{\mathbf{B}}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=1}^{n}\Big(\|\mathbf{x}_i(\mathbf{B}) - \mathbf{x}_i^\star\|_p + \|\boldsymbol{\lambda}_i(\mathbf{B}) - \boldsymbol{\lambda}_i^\star\|_p\Big) && \textit{minimize decision error} \\
\text{subject to} \quad & \mathbf{x}_i(\mathbf{B}) \cup \boldsymbol{\lambda}_i(\mathbf{B}) \in \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{argmin}} \quad \mathbf{x}^\top \mathbf{C}_i \mathbf{x} + \mathbf{c}_i^\top \mathbf{x} && \textit{FO for every } i \\
& \qquad\qquad\qquad\qquad \text{subject to} \quad \mathbf{A}_i \mathbf{x} \geqslant \mathbf{B}\boldsymbol{\varphi}_i : \boldsymbol{\lambda}
\end{aligned}$$

Data-driven IO for data generating models

- Consider that the right-hand side parameter $\mathbf{b}$ is variable
- Suppose it comes from a prediction model:

$$\widehat{\mathbf{b}} = \mathbf{B}\widehat{\boldsymbol{\varphi}}$$

  where $\widehat{\boldsymbol{\varphi}}$ is the feature vector, and $\mathbf{B}$ is the matrix of model weights
- Feature $\widehat{\boldsymbol{\varphi}}$ should correlate variable with $\widehat{\mathbf{b}}$ (e.g., as air temperature for loads)
- Given a history of observations $\{(\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star, \boldsymbol{\varphi}_i)\}_{i=1}^n$, find model weights $\mathbf{B}$

$$\begin{aligned}
\underset{\mathbf{B}}{\text{minimize}} \quad & \frac{1}{2}\sum_{i=1}^{n}\left(\|\mathbf{x}_i(\mathbf{B}) - \mathbf{x}_i^\star\|_p + \|\boldsymbol{\lambda}_i(\mathbf{B}) - \boldsymbol{\lambda}_i^\star\|_p\right) && \textit{minimize decision error} \\
\text{subject to} \quad & \mathbf{x}_i(\mathbf{B}) \cup \boldsymbol{\lambda}_i(\mathbf{B}) \in \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{argmin}} \quad \mathbf{x}^\top \mathbf{C}_i \mathbf{x} + \mathbf{c}_i^\top \mathbf{x} && \textit{FO for every } i \\
& \qquad\qquad\qquad\qquad \text{subject to} \quad \mathbf{A}_i \mathbf{x} \geqslant \mathbf{B}\boldsymbol{\varphi}_i : \boldsymbol{\lambda}
\end{aligned}$$

- n number of FO problems are coupled by common weights $\mathbf{B}$

Data-driven IO for data generating models

- Consider that the right-hand side parameter $\mathbf{b}$ is variable
- Suppose it comes from a prediction model:

$$\widehat{\mathbf{b}} = \mathbf{B}\widehat{\boldsymbol{\varphi}}$$

  where $\widehat{\boldsymbol{\varphi}}$ is the feature vector, and $\mathbf{B}$ is the matrix of model weights
- Feature $\widehat{\boldsymbol{\varphi}}$ should correlate variable with $\widehat{\mathbf{b}}$ (e.g., as air temperature for loads)
- Given a history of observations $\{(\mathbf{x}_i^\star, \boldsymbol{\lambda}_i^\star, \boldsymbol{\varphi}_i)\}_{i=1}^n$, find model weights $\mathbf{B}$

$$\begin{aligned}
&\underset{\mathbf{B}}{\text{minimize}} \quad \frac{1}{2}\sum_{i=1}^{n}\Big(\|\mathbf{x}_i(\mathbf{B}) - \mathbf{x}_i^\star\|_p + \|\boldsymbol{\lambda}_i(\mathbf{B}) - \boldsymbol{\lambda}_i^\star\|_p\Big) && \textit{minimize decision error} \\
&\text{subject to} \quad \mathbf{x}_i(\mathbf{B}) \cup \boldsymbol{\lambda}_i(\mathbf{B}) \in \underset{\mathbf{x},\boldsymbol{\lambda}}{\text{argmin}} \quad \mathbf{x}^\top \mathbf{C}_i \mathbf{x} + \mathbf{c}_i^\top \mathbf{x} && \textit{FO for every } i \\
&\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{subject to} \quad \mathbf{A}_i \mathbf{x} \geqslant \mathbf{B}\boldsymbol{\varphi}_i : \boldsymbol{\lambda}
\end{aligned}$$

- n number of FO problems are coupled by common weights $\mathbf{B}$
- **Exercise 1:** derive the gradient for SGD update
- **Exercise 2:** derive sensitivities $\frac{\partial \mathbf{x}(\mathbf{B})}{\partial \mathbf{B}}$ and $\frac{\partial \boldsymbol{\lambda}(\mathbf{B})}{\partial \mathbf{B}}$

Application: Unveiling ISO's wind power forecast model – I

- Consider an optimal power flow problem:

$$\begin{aligned}
\underset{\mathbf{p}}{\text{minimize}} \quad & \mathbf{p}^\top \mathbf{C}\mathbf{p} + \mathbf{c}^\top \mathbf{p} && && \textit{generator dispatch cost} \\
\text{subject to} \quad & \underline{\mathbf{p}} \leqslant \mathbf{p} \leqslant \overline{\mathbf{p}} && && \textit{generation limits} \\
& \mathbf{1}^\top(\mathbf{p} + \widehat{\mathbf{w}} - \mathbf{d}) = 0 && : \boldsymbol{\lambda}_b && \textit{power balance condition} \\
& |\mathbf{F}(\mathbf{p} + \widehat{\mathbf{w}} - \mathbf{d})| \leqslant \overline{\mathbf{f}} && : \boldsymbol{\lambda}_{\overline{\mathbf{f}}}, \boldsymbol{\lambda}_{\underline{\mathbf{f}}} && \textit{power flow limits}
\end{aligned}$$

  where $\widehat{\mathbf{w}}$ is the forecast of wind power generation
- **Goal:** unveil wind power forecast model of the system operator ...
- What information we would typically know in practice?

Application: Unveiling ISO's wind power forecast model – I

- Consider an optimal power flow problem:

$$
\begin{aligned}
\underset{\mathbf{p}}{\text{minimize}} \quad & \mathbf{p}^\top \mathbf{C}\mathbf{p} + \mathbf{c}^\top \mathbf{p} && && \textit{generator dispatch cost} \\
\text{subject to} \quad & \underline{\mathbf{p}} \leqslant \mathbf{p} \leqslant \overline{\mathbf{p}} && && \textit{generation limits} \\
& \mathbf{1}^\top(\mathbf{p} + \widehat{\mathbf{w}} - \mathbf{d}) = 0 && : \boldsymbol{\lambda}_b && \textit{power balance condition} \\
& |\mathbf{F}(\mathbf{p} + \widehat{\mathbf{w}} - \mathbf{d})| \leqslant \overline{\mathbf{f}} && : \boldsymbol{\lambda}_{\overline{\mathbf{f}}}, \boldsymbol{\lambda}_{\underline{\mathbf{f}}} && \textit{power flow limits}
\end{aligned}
$$

  where $\widehat{\mathbf{w}}$ is the forecast of wind power generation
- **Goal:** unveil wind power forecast model of the system operator ...
- What information we would typically know in practice?

- We observe public weather data $\boldsymbol{\varphi}$ and locational marginal prices $\boldsymbol{\pi}$

$$\boldsymbol{\pi} = \boldsymbol{\lambda}_b \cdot \mathbf{1} - \mathbf{F}^\top(\boldsymbol{\lambda}_{\overline{\mathbf{f}}} - \boldsymbol{\lambda}_{\underline{\mathbf{f}}})$$

Application: Unveiling ISO's wind power forecast model – I

- Consider an optimal power flow problem:

$$\begin{aligned}
\underset{\mathbf{p}}{\text{minimize}} \quad & \mathbf{p}^\top \mathbf{C}\mathbf{p} + \mathbf{c}^\top \mathbf{p} && && \textit{generator dispatch cost} \\
\text{subject to} \quad & \underline{\mathbf{p}} \leqslant \mathbf{p} \leqslant \overline{\mathbf{p}} && && \textit{generation limits} \\
& \mathbf{1}^\top(\mathbf{p} + \widehat{\mathbf{w}} - \mathbf{d}) = 0 && : \boldsymbol{\lambda}_b && \textit{power balance condition} \\
& |\mathbf{F}(\mathbf{p} + \widehat{\mathbf{w}} - \mathbf{d})| \leqslant \overline{\mathbf{f}} && : \boldsymbol{\lambda}_{\overline{\mathbf{f}}}, \boldsymbol{\lambda}_{\underline{\mathbf{f}}} && \textit{power flow limits}
\end{aligned}$$

  where $\widehat{\mathbf{w}}$ is the forecast of wind power generation
- **Goal:** unveil wind power forecast model of the system operator ...
- What information we would typically know in practice?

- We observe public weather data $\boldsymbol{\varphi}$ and locational marginal prices $\boldsymbol{\pi}$

$$\boldsymbol{\pi} = \boldsymbol{\lambda}_b \cdot \mathbf{1} - \mathbf{F}^\top(\boldsymbol{\lambda}_{\overline{\mathbf{f}}} - \boldsymbol{\lambda}_{\underline{\mathbf{f}}})$$

- **Goal (cont'd):** ... using historical LMPs and weather features

Application: Unveiling ISO's wind power forecast model – II

- Reformulate the problem in the FO form from above and take the dual

$$\begin{aligned} \underset{\mathbf{p}}{\text{minimize}} \quad & \frac{1}{2}\mathbf{p}^\top \mathbf{C}\mathbf{p} + \mathbf{c}^\top \mathbf{p} \\ \text{subject to} \quad & \mathbf{A}\mathbf{p} \geqslant \mathbf{b}(\widehat{\mathbf{w}}) \quad : \boldsymbol{\lambda} \end{aligned}$$

primal FO problem

$$\begin{aligned} \underset{\boldsymbol{\lambda}}{\text{maximize}} \quad & \mathbf{q}(\widehat{\mathbf{w}})^\top \boldsymbol{\lambda} - \boldsymbol{\lambda}^\top \mathbf{Q}\boldsymbol{\lambda} \\ \text{subject to} \quad & \boldsymbol{\lambda} \geqslant \mathbf{0} \end{aligned}$$

dual FO problem

where $\mathbf{q}(\widehat{\mathbf{w}}) = \mathbf{A}\mathbf{C}^{-1}\mathbf{c} + \mathbf{b}(\widehat{\mathbf{w}})$ and $\mathbf{Q} = \mathbf{A}\mathbf{C}^{-1}\mathbf{A}$

Application: Unveiling ISO's wind power forecast model – II

- Reformulate the problem in the FO form from above and take the dual

$$\begin{aligned} \underset{\mathbf{p}}{\text{minimize}} \quad & \frac{1}{2}\mathbf{p}^\top \mathbf{C}\mathbf{p} + \mathbf{c}^\top \mathbf{p} \\ \text{subject to} \quad & \mathbf{A}\mathbf{p} \geqslant \mathbf{b}(\widehat{\mathbf{w}}) \quad : \boldsymbol{\lambda} \end{aligned}$$

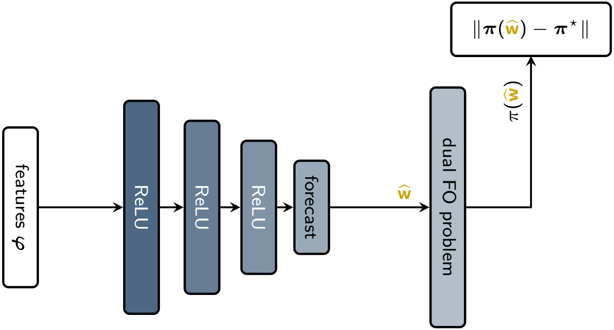
primal FO problem

$$\begin{aligned} \underset{\boldsymbol{\lambda}}{\text{maximize}} \quad & \mathbf{q}(\widehat{\mathbf{w}})^\top \boldsymbol{\lambda} - \boldsymbol{\lambda}^\top \mathbf{Q}\boldsymbol{\lambda} \\ \text{subject to} \quad & \boldsymbol{\lambda} \geqslant \mathbf{0} \end{aligned}$$

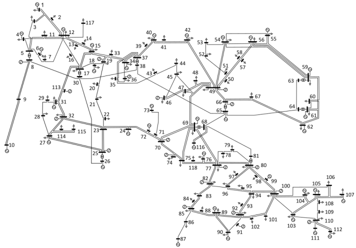
dual FO problem

where $\mathbf{q}(\widehat{\mathbf{w}}) = \mathbf{A}\mathbf{C}^{-1}\mathbf{c} + \mathbf{b}(\widehat{\mathbf{w}})$ and $\mathbf{Q} = \mathbf{A}\mathbf{C}^{-1}\mathbf{A}$

- We use a ReLU neural network to map features into predictions



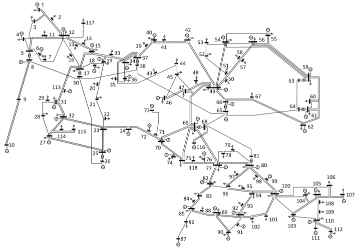
Application: Unveiling ISO's wind power forecast model – III



IEEE 118-bus system

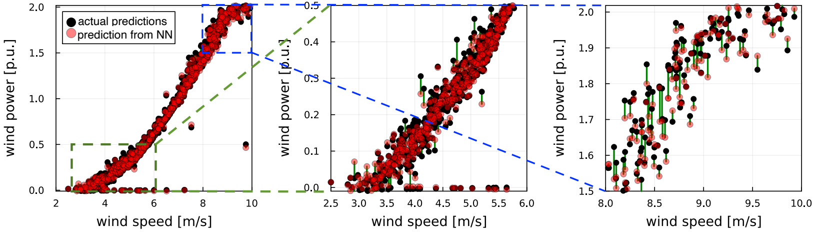
- One wind farm at bus 39
- $1,000$ scenarios with 118 LMPs, wind speed, direction, and peach blade angle
- Stochastic GD with 1,000 epochs

Application: Unveiling ISO's wind power forecast model – III



IEEE 118-bus system

- One wind farm at bus 39
- 1, 000 scenarios with 118 LMPs, wind speed, direction, and peach blade angle
- Stochastic GD with 1,000 epochs



Resources

- Chan, T. C., Mahmood, R., & Zhu, I. Y. (2025). Inverse optimization: Theory and applications. Operations Research, 73(2), 1046-1074. [**must read**]

- Mitridati, L., & Pinson, P. (2017). A Bayesian inference approach to unveil supply curves in electricity markets. IEEE Transactions on Power Systems, 33(3), 2610-2620.

- Birge, J. R., Hortaçsu, A., & Pavlin, J. M. (2017). Inverse optimization for the recovery of market structure from market outcomes: An application to the MISO electricity market. Operations Research, 65(4), 837-855.