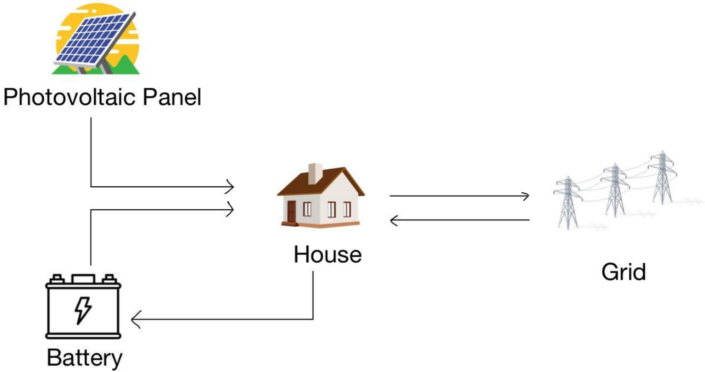
Computing Counterfactual Explanations of Linear Problems

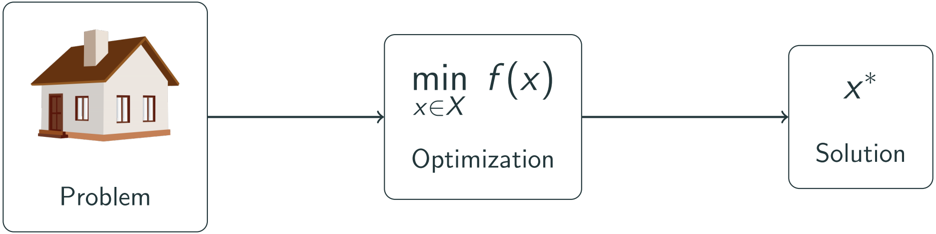
Henri Lefebvre, Martin Schmidt
Trier University (Germany), Department of Mathematics

Darmstadt, 2024

Understanding Energy Model Decisions



Understanding Energy Model Decisions



Understanding Energy Model Decisions

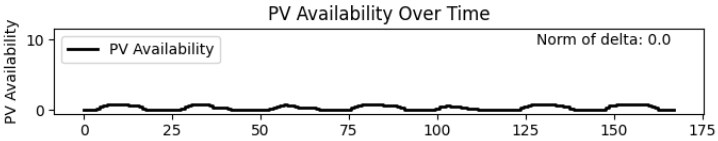
Problem

$$\min_{x \in X} \; f(x)$$

Optimization

$$x^*$$

Solution



Input Data

Outline

Outline

The Underlying Optimization Problem

We consider the linear optimization problem

$$\min_{y} \quad f^\top y \quad \text{s.t.} \quad \bar{D}y \geq b$$

The Underlying Optimization Problem

We consider the linear optimization problem

$$\min_{y} \quad f^\top y \quad \text{s.t.} \quad \bar{D} y \geq b$$

Its modified counterpart reads

$$\min_{y} \quad f^\top y \quad \text{s.t.} \quad D(x) y \geq b$$

with $d_{ij}(x) = \bar{d}_{ij} + \tilde{d}_{ij}^\top x$ for some $x \in X$

The Underlying Optimization Problem

We consider the linear optimization problem

$$\min_{y} \quad f^\top y \quad \text{s.t.} \quad \bar{D}y \geq b$$

Its modified counterpart reads

$$\min_{y} \quad f^\top y \quad \text{s.t.} \quad D(x)y \geq b$$

with $d_{ij}(x) = \bar{d}_{ij} + \tilde{d}_{ij}^\top x$ for some $x \in X$

Goal: Find an x so that an optimal point y^* is in the desired space Y

Problem Formulation

$$
\begin{aligned}
\inf_{x,y} \quad & f(x) \\
\text{s.t.} \quad & x \in X \\
& y \in Y \\
& y \in \arg\min_{\bar{y}} \left\{ f^\top \bar{y} \;:\; D(x)\bar{y} \geq b \right\}
\end{aligned}
$$

Strong-Duality Reformulation

Bilevel Constraint

$$y \in \arg\min_{\bar{y}} \left\{ f^\top \bar{y} \;:\; D(x)\bar{y} \geq b \right\}$$

Strong-Duality Reformulation

Bilevel Constraint

$$y \in \arg\min_{\bar{y}} \left\{ f^\top \bar{y} \ : \ D(x)\bar{y} \geq b \right\}$$

$\downarrow$

Single-Level Constraint

There exists λ s.t.

$$D(x)y \geq b$$
$$D(x)^\top \lambda = f, \quad \lambda \geq 0$$
$$f^\top y \leq b^\top \lambda$$

Strong-Duality Reformulation

Bilevel Constraint

$$y \in \arg\min_{\bar{y}} \left\{ f^\top \bar{y} \ : \ D(x)\bar{y} \geq b \right\}$$

$\downarrow$

Single-Level Constraint

There exists λ s.t.

$$D(x)y \geq b$$
$$D(x)^\top \lambda = f, \quad \lambda \geq 0$$
$$f^\top y \leq b^\top \lambda$$

$$\begin{aligned} \inf_{x,y} \quad & f(x) \\ \text{s.t.} \quad & x \in X, y \in Y \\ & y \in \arg\min_{\bar{y}} \left\{ f^\top \bar{y} \ : \ D(x)\bar{y} \geq b \right\} \end{aligned}$$

Strong-Duality Reformulation

Bilevel Constraint

$$y \in \arg\min_{\bar{y}} \left\{ f^\top \bar{y} \ : \ D(x)\bar{y} \geq b \right\}$$

$$\inf_{x,y} \quad f(x)$$
$$\text{s.t.} \quad x \in X, y \in Y$$
$$y \in \arg\min_{\bar{y}} \left\{ f^\top \bar{y} \ : \ D(x)\bar{y} \geq b \right\}$$

↓

Single-Level Constraint

There exists λ s.t.

$$D(x)y \geq b$$
$$D(x)^\top \lambda = f, \quad \lambda \geq 0$$
$$f^\top y \leq b^\top \lambda$$

↓

$$\inf_{x,y} \quad f(x)$$
$$\text{s.t.} \quad x \in X, y \in Y$$
$$D(x)y \geq b$$
$$D(x)^\top \lambda = f, \quad \lambda \geq 0$$
$$f^\top y \leq b^\top \lambda$$

Outline

A General Optimization Problem

$$
\begin{aligned}
\min_{x, y} \quad & F(x, y) \\
\text{s.t.} \quad & x \in \mathcal{X}, \quad y \in \mathcal{Y}, \\
& g(x, y) \leq 0.
\end{aligned}
$$

A General Optimization Problem

$$\begin{aligned} \min_{x,y} \quad & F(x, y) \\ \text{s.t.} \quad & x \in \mathcal{X}, \quad y \in \mathcal{Y}, \\ & g(x, y) \leq 0. \end{aligned}$$

Assumption: Solving over x for fixed y is easy, and the other way around.

Alternating Direction Method

Given: Initial values $(x^0, y^0) \in \mathcal{X} \times \mathcal{Y}$.

for $i = 0, 1, \ldots$ **do**

Choose $x^{i+1} \in \arg\min_{x} \left\{ F(x, y^i) : g(x, y^i) \leq 0, x \in \mathcal{X} \right\}$.

Choose $y^{i+1} \in \arg\min_{y} \left\{ F(x^{i+1}, y) : g(x^{i+1}, y) \leq 0, y \in \mathcal{Y} \right\}$.

end for

Alternating Direction Method

Given: Initial values $(x^0, y^0) \in \mathcal{X} \times \mathcal{Y}$.

for $i = 0, 1, \ldots$ **do**

Choose $x^{i+1} \in \arg\min_x \left\{ F(x, y^i) : g(x, y^i) \leq 0, x \in \mathcal{X} \right\}$.

Choose $y^{i+1} \in \arg\min_y \left\{ F(x^{i+1}, y) : g(x^{i+1}, y) \leq 0, y \in \mathcal{Y} \right\}$.

end for

Good (from a theoretical viewpoint): Convergence is well understood

Alternating Direction Method

Given: Initial values $(x^0, y^0) \in \mathcal{X} \times \mathcal{Y}$.

for $i = 0, 1, \ldots$ **do**

Choose $x^{i+1} \in \arg\min_{x} \left\{ F(x, y^i) : g(x, y^i) \leq 0, x \in \mathcal{X} \right\}$.

Choose $y^{i+1} \in \arg\min_{y} \left\{ F(x^{i+1}, y) : g(x^{i+1}, y) \leq 0, y \in \mathcal{Y} \right\}$.

end for

Good (from a theoretical viewpoint): Convergence is well understood

Issues (from a practical viewpoint):

1. Sub-problems may be infeasible because of the coupling constraints
2. Poor performance

A Penalized Problem

For some penalty parameter $\mu \in \mathbb{R}^m_{>0}$,

$$\begin{aligned} \min_{x,y} \quad & F(x,y) \\ \text{s.t.} \quad & x \in \mathcal{X}, \quad y \in \mathcal{Y} \\ & g(x,y) \leq 0 \end{aligned}$$

$$\downarrow$$

$$\begin{aligned} \min_{x,y} \quad & F(x,y) + \sum_{i=1}^{m} \mu_i [g_i(x,y)]^+ \\ \text{s.t.} \quad & x \in \mathcal{X}, \quad y \in \mathcal{Y} \end{aligned}$$

$[u]^+ = \max\{0, u\}$

Penalty Alternating Direction Method

Given: Initial values $(x^{0,0}, y^{0,0}) \in \mathcal{X} \times \mathcal{Y}$ and $\mu^0 \in \mathbb{R}^r_{\geq 0}$.

for $j = 0, 1, \ldots$ **do**

Set $i \leftarrow 0$

while $(x^{i,j}, y^{i,j})$ is not a partial minimizer of the penalized problem **do**

Choose $x^{i+1} \in \arg\min_x \left\{\phi(x, y^i; \mu) : x \in \mathcal{X}\right\}$

Choose $y^{i+1} \in \arg\min_y \left\{\phi(x^{i+1}, y; \mu) : y \in \mathcal{Y}\right\}$

Set $i \leftarrow i + 1$

end while

Choose new penalty parameters $\mu^{j+1} \geq \mu^j$.

end for

Penalty Alternating Direction Method

Given: Initial values $(x^{0,0}, y^{0,0}) \in \mathcal{X} \times \mathcal{Y}$ and $\mu^0 \in \mathbb{R}^r_{\geq 0}$.

for $j = 0, 1, \dots$ **do**

Set $i \leftarrow 0$

while $(x^{i,j}, y^{i,j})$ is not a partial minimizer of the penalized problem **do**

Choose $x^{i+1} \in \arg\min_{x} \{\phi(x, y^i; \mu) : x \in \mathcal{X}\}$

Choose $y^{i+1} \in \arg\min_{y} \{\phi(x^{i+1}, y; \mu) : y \in \mathcal{Y}\}$

Set $i \leftarrow i + 1$

end while

Choose new penalty parameters $\mu^{j+1} \geq \mu^j$.

end for

Good: Convergence is well understood, e.g., Geißler et al. (2017).

Outline

Applying the PADM to the Single-Level Reformulation

Single-Level Reformulation

$$
\begin{aligned}
\inf_{x,y} \quad & f(x) \\
\text{s.t.} \quad & x \in X, \quad y \in Y \\
& D(x)y \geq b \\
& D(x)^\top \lambda = f, \quad \lambda \geq 0 \\
& f^\top y \leq b^\top \lambda
\end{aligned}
$$

Applying the PADM to the Single-Level Reformulation

Single-Level Reformulation

$$
\begin{aligned}
\inf_{x,y} \quad & f(x) \\
\text{s.t.} \quad & x \in X, \quad y \in Y \\
& D(x)y \geq b \\
& D(x)^\top \lambda = f, \quad \lambda \geq 0 \\
& f^\top y \leq b^\top \lambda
\end{aligned}
$$

Penalized Problem

$$
\begin{aligned}
\min_{x,y,\lambda} \quad & f(x) + \sum_{i=1}^{m} \rho_i \left[b_i - d_{i\cdot}(x)y\right]^+ + \sum_{j=1}^{n} \mu_j \left|c_j - d_{\cdot j}(x)^\top \lambda\right| \\
\text{s.t.} \quad & x \in X, \quad y \in Y, \quad \lambda \geq 0, \quad f^\top y \leq b^\top \lambda
\end{aligned}
$$

The Sub-problems

1. Try to find feasible primal-dual point given $\hat{x}$ by solving

$$\min_{y,\lambda} \quad \sum_{i=1}^{m} \rho_i \left[b_i - d_{i\cdot}(\hat{x})y\right]^+ + \sum_{j=1}^{n} \mu_j \left|c_j - d_{\cdot j}(\hat{x})^\top \lambda\right|$$
$$\text{s.t.} \quad y \in Y, \quad \lambda \geq 0, \quad f^\top y \leq b^\top \lambda$$

The Sub-problems

1. Try to find feasible primal-dual point given $\hat{x}$ by solving

$$\min_{y,\lambda} \quad \sum_{i=1}^{m} \rho_i \left[b_i - d_{i\cdot}(\hat{x})y\right]^+ + \sum_{j=1}^{n} \mu_j \left|c_j - d_{\cdot j}(\hat{x})^\top \lambda\right|$$

$$\text{s.t.} \quad y \in Y, \quad \lambda \geq 0, \quad f^\top y \leq b^\top \lambda$$

2. Try to repair infeasibilities of $(\hat{y}, \hat{\lambda})$ by choosing a new x

$$\min_{x} \quad f(x) + \sum_{i=1}^{m} \rho_i \left[b_i - d_{i\cdot}(x)\hat{y}\right]^+ + \sum_{j=1}^{n} \mu_j \left|c_j - d_{\cdot j}(x)^\top \hat{\lambda}\right|$$

$$\text{s.t.} \quad x \in X$$

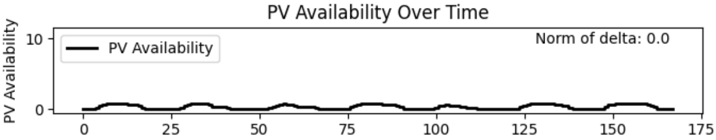
Outline

Outline

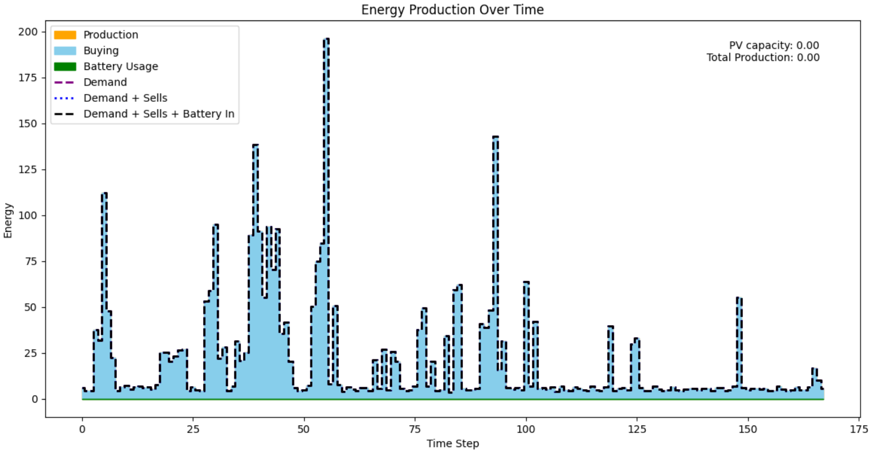
One Example

We consider a given week with

1. costs for buying electricity
2. price for selling electricity surplus
3. house demands
4. cost for buying photovolatic panels
5. historical data on PV availability:



Status Quo: Only Buying From the Grid



The Question

Keeping the current prices.
What should the PV availability be like so that one invests in PV panels?

The Question

> Keeping the current prices.
> What should the PV availability be like so that one invests in PV panels?

More explicitly: we ask for producing 1000 kWh during the week.

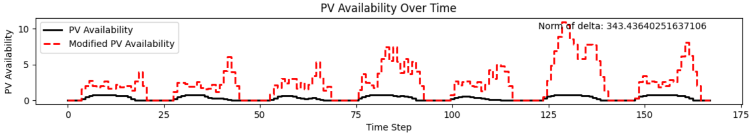
The Question

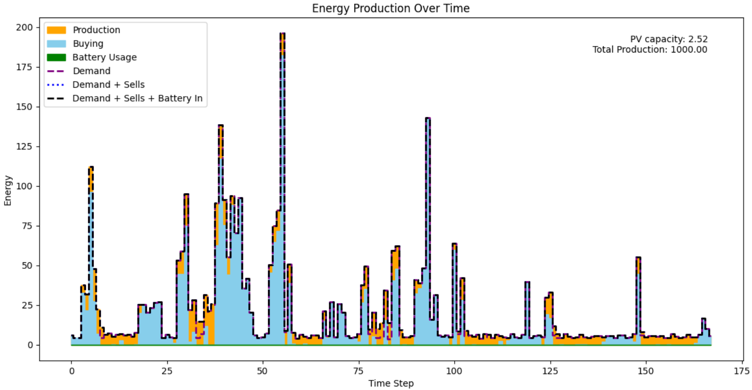
Keeping the current prices.
What should the PV availability be like so that one invests in PV panels?

More explicitly: we ask for producing 1000 kWh during the week.

Answer...

Modified Solution





Outline

Setting

Instances based on the NETLIB library, adapted by Kurtz et al. (2024)

Type	Category	Intervals
# Variables	small	$0 \leq n \leq 534$
	medium	$534 \leq n \leq 2167$
	large	$2167 \leq n \leq 22275$
# Constraints	small	$0 \leq m \leq 351$
	medium	$351 \leq m \leq 906$
	large	$906 \leq m \leq 16675$

Setting

Instances based on the NETLIB library, adapted by Kurtz et al. (2024)

Type	Category	Intervals
# Variables	small	$0 \leq n \leq 534$
	medium	$534 \leq n \leq 2167$
	large	$2167 \leq n \leq 22275$
# Constraints	small	$0 \leq m \leq 351$
	medium	$351 \leq m \leq 906$
	large	$906 \leq m \leq 16675$

- Randomly selects 1, 5, or 10 columns which are mutable

Setting

Instances based on the NETLIB library, adapted by Kurtz et al. (2024)

Type	Category	Intervals
# Variables	small	$0 \leq n \leq 534$
	medium	$534 \leq n \leq 2167$
	large	$2167 \leq n \leq 22275$
# Constraints	small	$0 \leq m \leq 351$
	medium	$351 \leq m \leq 906$
	large	$906 \leq m \leq 16675$

- Randomly selects 1, 5, or 10 columns which are mutable
- A total of 5 760 instances

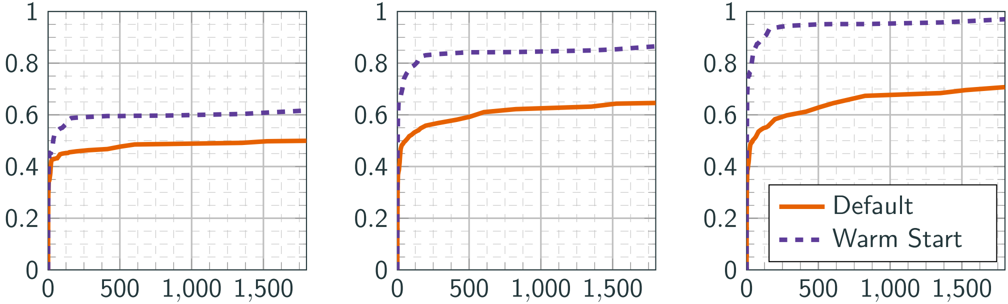
Setting

Instances based on the NETLIB library, adapted by Kurtz et al. (2024)

Type	Category	Intervals
# Variables	small	$0 \leq n \leq 534$
	medium	$534 \leq n \leq 2167$
	large	$2167 \leq n \leq 22275$
# Constraints	small	$0 \leq m \leq 351$
	medium	$351 \leq m \leq 906$
	large	$906 \leq m \leq 16675$

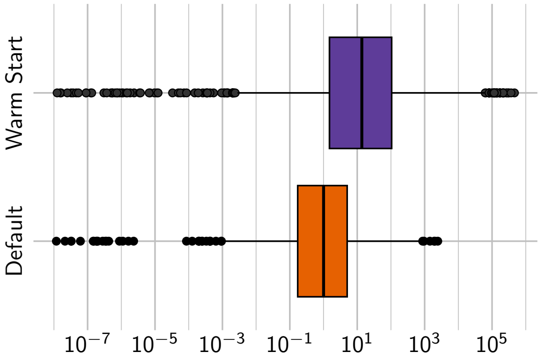
- Randomly selects 1, 5, or 10 columns which are mutable
- A total of 5 760 instances
- Not all instances are feasible ($\sim$ 55%)

ECDF of computation times for $f = \|\cdot\|_1$



(a) 1 mutable column **(b)** 5 mutable columns **(c)** 10 mutable columns

Solution Quality



Outline

Conclusion

What We Have Done:

1. Computing counterfactual explanations of linear problems is challenging
2. We derive a single-level reformulation which we heuristically solve by the PADM
3. We can characterize the computed solutions (stationary points of the single-level reformulation)

What to Do Then:

1. Polish the preprint and submit
2. Can we do the same for strong counterfactual explanations
3. Discuss next project steps