Strategyproof Linear Regression in High Dimensions

Yiling Chen, Chara Podimata, Ariel D. Procaccia and Nisarg Shah

Harvard University, Harvard University, Carnegie Mellon University, University of Toronto

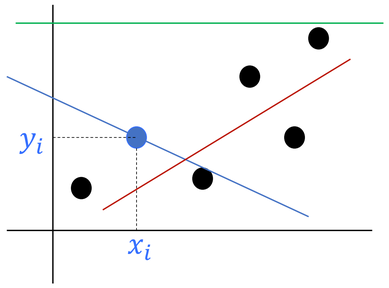
Strategic Noise – Why Different

Strength of "opponent" →

Stochastic Noise	Strategic Noise	Adversarial Noise
(errors ~ Distribution)	(**any** error, s.t., loss < gain)	(**any** error)

Model

- n agents, each controls **one** datapoint: $(x_i, y_i) \in \mathbb{R}^{d+1}$
- x_i: **public** information, y_i: **private** (manipulable)
- Agents are **strategic**:
  - Can report: $(x_i, \tilde{y_i}) \in \mathbb{R}^{d+1}$ s.t., $\tilde{y_i} \neq y_i$
  - Single Peaked Preferences: prefer outcome y_i



Any line (β_1, β_0) *closer to*:
$y_i = \beta_1 x_i + \beta_0$
is preferred by agent (x_i, y_i)
[e.g., blue ≻ red ≻ green]

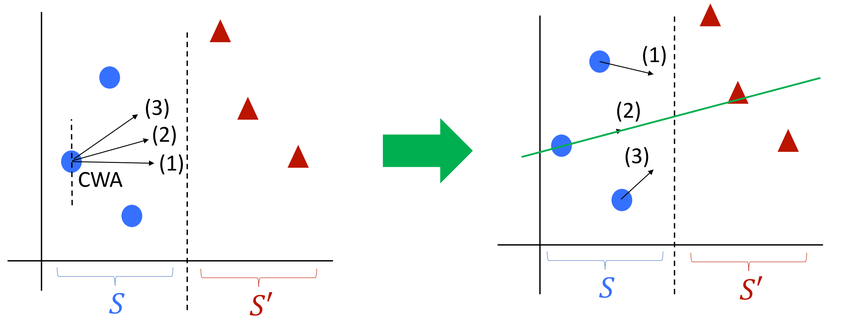
Our Goal. Construct Linear Regression mechanism $M^{x}(\tilde{y})$:
- Learn relationship between $\boldsymbol{x}, \boldsymbol{y}$
- Induce **strategyproofness (SP)** (prefer truthtelling from misreporting, irrespective of reports of others) without **monetary** incentives

Main Result

New Mechanism for Group-Strategyproof High Dimensional Linear Regression (using the Ham Sandwich Theorem)

Example of a SP Linear Regression

(corrected) CRM Family [Perote & Perote-Peña, 2004]



1) Split dataset into S, S' s.t.:
$S = S'$ or S, S': separable
2) From each point in S, compute CWA for points in S'

3) Final line:
median-of-median CWAs
- Only defined for 2D, **not** generalizable to higher dimensions

Generalized Resistant Lines (GRL)

Definition $[(S, S', k, k') - \text{GRL}]$. Choose S, S' separable and $k \in [|S|], k' \in [|S'|]$. Output line (β_1, β_0) s.t.:
$\min_{i \in S}^{k}(y_i - \beta_1 x_i - \beta_0) = \min_{j \in S'}^{k'}(y_j - \beta_1 x_j - \beta_0) = 0$

E.g., previous example was a $(S, S', 2, 2)$-GRL mechanism

In fact, CRM ⊆ GRL!

GRL mechanisms generalize to higher dimensions.
Problems:
1) Definition of **separability** in higher dimensions
2) Uniqueness of solution

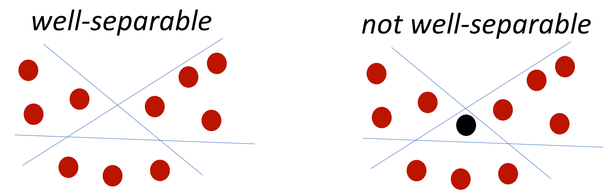
Ham Sandwich Theorem

Theorem (Stone & Tukey, 1942). Given k continuous measures $\mu_1, \dots, \mu_k$ *on* $\mathbb{R}^k$, ∃ hyperplane, H: $\mu_i(H^+) = \frac{1}{2}, \forall i \in [k]$.

- *Discrete Variant* (still bisecting) due to [Elton & Hill, 2011]
- *Unique Variant* due to [Steiger and Zhao, 2010]
  → close to what we need

Generalized Resistant Hyperplanes (GRH)

Definition. Choose $S_1, \dots, S_{d+1}$ **publicly separable** and $k_1 \in [|S_1|], \dots, k_{d+1} \in [|S_{d+1}|]$. Output line $(\beta_{d+1}, \dots, \beta_0)$:
$\min_{i \in S_t}^{k_t}(y_i - \beta^T \boldsymbol{x}_i) = 0, \forall t \in [d+1]$



Theorem. GRH mechanisms yield **unique solution** and are **group-strategyproof**.

Proof Idea.
- Any coalition creates a new hyperplane.
- Uniqueness of GRH for given $k_1, \dots, k_{d+1}$ → new hyperplane either does not exist, or some agent is not rational

Efficiency of SP Linear Regressors

When **no** strategic considerations: Ordinary Least Squares (OLS) is most popular, **but**

OLS is **not** strategyproof!

Any SP Linear Regressor has Residual Sum of Squares (RSS) error at least **twice** as large as RSS(OLS)!

Open Questions

1) What about **consistency**?
2) SP linear regressors for **other types of** agent **incentives**?
3) Constructive characterization of **all** SP linear regressors.

Overarching Goal
To build a **Theory of Incentives** for ML algorithms.