Geometry of Relaxed Fair Regression : A Unified Framework for Aware and Unaware Settings

Fairness-accuracy trade-offs are a central concern in the deployment of fairness-aware machine learning methods. When sensitive attributes are unavailable at inference time–the so called unawareness setting, principled methods for obtaining accurate predictions under relaxed fairness constraints are largely missing. In this work, we address this gap by formulating regression under a demographic parity penalty as an optimal transport problem. Our framework, unifies both the *aware* and *unaware* settings and characterizes optimal prediction functions via optimal transport maps, under both squared Wasserstein-2 and Total Variation penalties. These results reveal that the choice of penalty reflects fundamentally different fairness philosophies : the Wasserstein penalty induces a smooth, population-wide compromise, while Total Variation enforces parity for a subset of individuals. Building on these theoretical characterizations, we propose FAUST, an algorithm that is simple to implement, computationally efficient, and consistently matches or outperforms state-of-the-art baselines on real-world benchmarks.