Learning Under Adversarial and Interventional Shifts

Harvineet Singh[*]
New York University

Finale Doshi-Velez
Harvard University

Himabindu Lakkaraju
Harvard University

Abstract

Machine learning models are often trained on data from one distribution and deployed on others. So it becomes important to design models that are robust to distribution shifts. Robustness is typically considered either against adversarial or interventional shifts in data. However, neither of these are ideal standalone. While adversarial shifts are often unrealistic, interventional shifts can result in conservative models. In this work, we address these shortcomings and propose a new formulation for designing models that are robust to a set of distribution shifts that are at the intersection of adversarial and interventional shifts. We employ the distributionally-robust optimization framework to optimize the resulting objective both in supervised as well as reinforcement learning settings. We also carry out synthetic experiments to demonstrate the promise of the proposed framework.

1 Introduction

To improve trust in machine learning models prior to deployment, one of the desired properties is that the models are robust to distribution shifts. This is particularly important for deployments in consequential decision-making such as in healthcare and criminal justice. Guaranteeing good predictive performance on (unseen) test distributions requires departing from choosing the best model for the train distribution. Two broad principles have been proposed for learning robust models *without* access to samples from test distribution – adversarial training (e.g. [20, 22, 10]) and causal invariance (e.g. [24, 1, 13]). Interestingly, recent work has shown correspondence between the two within the framework of *distributionally-robust optimization* (DRO) [2, 16]. Our work is motivated by this unification and aims to address some limitations of the current approaches to robust learning.

In DRO, the goal is to optimize a *worst-case* loss over a set of distributions defined to be *close* in some metric to a nominal distribution (typically, the train distribution). The set is referred to as the uncertainty set as it encodes our uncertainty about the test distribution. Under a large class of divergence metrics for defining such sets, one can give rigorous generalization guarantees owing to the worst-case nature of the optimization [4, 23]. At the same time, if the uncertainty set is too large, the DRO solution can be too pessimistic. That is, minimizing for the worst-case loss will result in high loss on the test distributions we actually encounter. For example, considering perturbations in images like adversarial noise or rotations can result in large sets which do not correspond to 'realistic' shifts [25]. Depending on the type of models and available sample sizes, robust learning can significantly degrade performance on train distributions [15]. Hence, the balance between utility and robustness of the models depends critically on the definition of the uncertainty sets. Prior work considers description of such sets using either divergence metrics or interventions on causal models. We consider a combination of such descriptions which allows expressing more realistic train-test shifts and learning less conservative models that are robust to such shifts.

[*] Work done as a summer fellow at Center for Research on Computation and Society, Harvard University. Correspondence at hs3673@nyu.edu.

Causal Discovery & Causality-Inspired Machine Learning Workshop at Neural Information Processing Systems, 2020.

Our contribution. We describe a natural way of specifying uncertainty sets defined using bounded interventions on causal models of the domain. We give an efficient procedure for solving the new DRO problem for such sets under certain conditions. The procedure is applied to problems from supervised and reinforcement learning on synthetic datasets. In summary, we provide a novel perspective to the problem of robust learning that bridges existing adversarial learning and causal inference methods.

2 Preliminaries

Consider a random variable V that will denote all observed variables in the domain. Throughout the text, we will denote train and test distributions of V by P and Q, respectively (using the same notation for their densities). Assume that the test distribution lies in the uncertainty set $\mathcal{U}_P$ defined with respect to the train distribution P. In DRO, our objective is to find a decision variable θ (e.g. regression parameters) from some set Θ that minimizes the robust loss,

$$\text{(DRO)} \qquad \theta^* \in \arg\min_{\theta \in \Theta} \left\{ \mathcal{R}(\theta, \mathcal{U}_P) := \sup_{Q \in\, \mathcal{U}_P} \mathbb{E}_{V \sim Q}[\ell(\theta, V)] \right\}, \tag{1}$$

where $\ell(\theta, V)$ is a given loss function. This problem can be seen as a game between the the min player (modeller) trying to minimize expected loss for the test distribution and the max player (adversary) trying to find the worst-case test distribution from the given uncertainty set.

We will first consider the supervised learning problem where the variable $V{:=}(X, Y)$ denotes i.i.d. features and outcome pair, and decision variable θ corresponds to the prediction function. In Section 3.2, we will consider an offline reinforcement learning problem and will introduce notation there.

2.1 Different Uncertainty Sets

Prior work has considered two ways of defining uncertainty sets – divergence-based and intervention-based – that encode knowledge of how train and test distributions differ. Next, we describe the two types of sets, followed by our proposal that combines them.

Divergence-based sets Uncertainty set $\mathcal{U}_P$ can be defined as all distributions which lie in a δ-ball around the train distribution defined with respect to a divergence metric. For a metric $D[\cdot\|\cdot]$, e.g f-divergence, the uncertainty set is defined as,

$$\mathcal{U}_P^{\text{div}} := \{Q \ll P \text{ s.t. } D[Q\|P] \leq \delta\} \tag{2}$$

Here, $Q \ll P$ denotes absolute continuity i.e. $P(V) = 0$ implies $Q(V) = 0$. Extensive literature exists on efficient computational procedures to learn under such shifts with favourable statistical properties [4, 18]. Other examples of sets are based on Wasserstein metric [7] and MMD [23].

However, the optimization problem (1) is known to result in degenerate solutions in some cases when uncertainty sets are large [9]. Thus, recent work has explored multiple ways to restrict the sets, mainly by asserting the existence of factors in the joint distribution $P(V)$ that remain the same across train and test [9, 21] e.g. assuming that there exists features $Z \subset V$ s.t. $P(V \mid Z) = Q(V \mid Z), \forall Q \in \mathcal{U}_P$. We aim to generalize the process of describing (or inferring) such invariant factors.

Intervention-based sets Causal knowledge can be exploited to guarantee distributional-robustness, as shown in recent works [16, 24, 11, 17, 14]. In particular, using causal parents of the outcome Y as features yields a prediction function that is robust to arbitrary interventions on features X. Further, Rojas-Carulla et al. [16, Theorem 4] shows a correspondence between the causal solution (using causal parents of Y) and the DRO solution with a particular uncertainty set. It shows that the causal solution minimizes the worst-case risk across distributions resulting from arbitrary (stochastic) interventions on X, i.e. the corresponding uncertainty set can be written as,

$$\mathcal{U}_P^{\text{int}} := \{Q \ll P \text{ s.t. } Q = P(V \mid do(X \sim \nu(X))), \nu \text{ is any distribution}\}, \tag{3}$$

where $do(\cdot)$ is the do-intervention [12]. Here, the $do(\cdot)$ notation is used for a *soft* intervention that artificially manipulates the underlying mechanism generating X [3], changing the distribution of X from $P(X)$ to a new distribution $\nu(X)$. An important distinction from divergence-based sets is allowing for *arbitrary* shifts in $P(X)$ instead of a bounded shift in some metric. The solution to (1) with the set $\mathcal{U}_P^{\text{int}}$ is to use the parents of Y, say $pa(Y)$, as features for learning the robust predictor [16]. That is, for the squared loss, the DRO solution is the Bayes predictor $\mathbb{E}[Y|pa(Y)]$.[2]

One drawback of causal approaches is that they can result in functions with high loss as they require robustness to arbitrary shifts, hence, leading to large uncertainty sets. When we expect the train and test distributions to not vary much, then the robust loss under the intervention-based set gives an overestimate of the plausible loss that we will observe at test time. Further, the advantage of the worst-case guarantee for the causal solution is limited to cases where we can find the Bayes predictor in the considered function class with reasonable sample sizes. As we will illustrate in the experiments, model misspecification can invalidate this approach's attractive guarantees to arbitrary shifts. Next we introduce our approach for defining the uncertainty sets that consists of realistic shifts, and thus, mitigates the drawbacks of the above two approaches.

3 Our Approach – Intersection of Divergence and Intervention-based Sets

Our key idea is to leverage both the approaches such that we can specify realistic shifts using interventions (as opposed to divergences) with bounded magnitude (as opposed to arbitrary interventions). Defining the uncertainty set this way may help the modeller to achieve a better trade-off between utility and robustness. Next, we describe *selection diagrams* used to specify the proposed sets.

Representing plausible shifts Assume we are given the causal graph for our domain along with annotations for plausible shifts. One such representation is selection diagrams which annotates the graph with additional nodes pointing to variables whose mechanisms may differ across train and test environments [13]. Selection diagrams have been used for constructing robust predictors [24, 11]. An example is given in Figure 1 with two features $X := (Z, T)$, outcome Y, and shift indicator E.

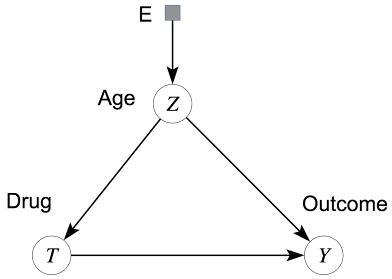


Figure 1: **Example causal graph.** Drugs are prescribed based on age and have different outcomes moderated by age. A domain expert posits that age distribution might change across hospitals. Selection node E (for environment) specifies this plausible distribution shift, denoting that $P(Z)$ may change from one environment to another.

Our proposal for the uncertainty set at the intersection of the two approaches is,

$$\mathcal{U}_P^{\text{div}\cap\text{int}} := \{Q \ll P \text{ s.t. } Q = P(V \mid do(Z \sim \nu(Z))), D[\nu(Z) \| P(Z)] \leq \delta, Z := ch(E)\}. \tag{4}$$

where Z are the children of the selection node, $ch(E)$. In words, the proposed uncertainty set contains all distributions resulting from interventions on selection nodes which are bounded in some metric. As we can observe, this definition limits the arbitrary shifts in (3) by requiring bounded shifts. Thus, we can see (4) as the intersection of (2) and (3), which reduces the size of the uncertainty set. The notation assumes no other variable causes Z for conciseness. If Z has parent nodes, then the divergence $D[\cdot\|\cdot]$ can be defined for the conditional distribution of $Z|pa(Z)$ with a technical condition for its existence. We note that similar sets based on bounded shift interventions have been proposed by Rothenhäusler et al. [17] for the supervised learning case. But, these assume a linear Gaussian model for all variables and consider only shifts that change the mean of the variables. In contrast, the proposed sets are defined without any parametric assumptions on variables or shifts.

Solving the DRO problem The optimization problem (1) now includes minimizing the worst-case loss, $\mathcal{R}(\theta, \mathcal{U}_P^{\text{div}\cap\text{int}})$, which is a function of the *interventional* distributions. Thus, identification of the causal estimand, worst-case loss in this case, requires making assumptions about the augmented causal graph. To make the problem tractable, we make the following assumptions.

Assumption 1. *For the causal graph $\mathcal{G}$ with the selection node E, assume the following holds.*

[2] The result requires that the mechanism between X and Y does not change due to the intervention on X [2].

(a) $\mathcal{G}$ is a directed acyclic graph with no unmeasured confounding, i.e. $P(V) = \prod_{O \in V} P(O|pa(O))$,

(b) $ch(E)$ have no parents except E, i.e. shifts only occur in variables with no parents.

While Assumption 1(a) is a statement about thoroughness of the measured variables and is frequently made in domains such as epidemiology [8], Assumption 1(b) stems from the tractability of solving the minmax problem in (1). Nonetheless, such conditions still cover many important shifts as demonstrated in experiments. The following result allows us to efficiently solve the DRO problem.

Proposition 1. *Given Assumption 1 holds, the uncertainty set for the intersection (4) is equivalent to,*

$$\mathcal{U}_P^{\mathit{div} \cap \mathit{int}} = \{Q \ll P \text{ s.t. } Q = \nu(Z)P(V \setminus Z \mid Z), D[\nu(Z) \| P(Z)] \leq \delta, Z := ch(E)\}. \tag{5}$$

This means that the proposed uncertainty set consists of distributions with bounded shifts in the marginals of some feature set Z. The proof is a straight-forward application of factorization from Assumption 1(a), followed by replacing factors $P(Z)$ with the intervened distribution $\nu(Z)$ by the definition of the *do*-intervention. We observe that the expression (5) does not contain any interventional terms. Thus, the causal estimand is identified using observations from the train distribution alone.

The remaining issue is that of solving the minmax problem with $\mathcal{U}_P^{\text{div}\cap\text{int}}$. Recent work by Duchi et al. [5] introduces a procedure for solving DRO under marginal shifts. The uncertainty set that they consider is equivalent to $\mathcal{U}_P^{\text{div}\cap\text{int}}$ in (5) with a particular divergence metric D. Thus, we can leverage their approach to solve our minmax problem. For completeness, we first describe the approach of Duchi et al. [5] in context of the supervised learning problem, as originally proposed. In Section 3.2, we present a novel extension to the problem of off-policy evaluation.

3.1 Robust Supervised Learning – Existing Work by Duchi et al. [5]

Consider the robust supervised learning problem for a dataset $\{V_i := (X_i, Y_i)\}_i$ containing n samples. The test distribution changes due to shifts in marginals of features $Z \subseteq X$ defined in terms of *subpopulation shifts* [5]. That is, the uncertainty set is built such that the train population contains at least δ_0 proportion of the test population.

$$\mathcal{U}_P^{\text{sub}} := \{Q(Z)P(V \setminus Z|Z) \text{ s.t. } P(Z) = \delta Q(Z) + (1-\delta)Q'(Z), \delta_0 \leq \delta \leq 1\}, \tag{6}$$

where $Q'(Z)$ is any distribution and $\delta_0 \in (0, 1]$ determines the minimum size of the subpopulation shared between train and test. We note that this set is an instance of (5) with a particular divergence metric, which can be defined in terms of the likelihood ratio $Q(Z)/P(Z)$. Using convex duality arguments [18], one can write the worst-case loss alternatively as,

$$\mathcal{R}(\theta, \mathcal{U}_P^{\text{sub}}) = \sup_{Q \in \mathcal{U}_P^{\text{sub}}} \mathbb{E}_{Z \sim Q(Z)} \mathbb{E}_{P(V|Z)}[\ell(\theta, V)|Z] = \inf_{\eta \in \mathbf{R}} \frac{1}{\delta_0} \mathbb{E}_P\left[(\mathbb{E}_P\left[\ell(\theta, V)|Z\right] - \eta)_+\right] + \eta,$$

where $(\cdot)_+ = \max(\cdot, 0)$. Note that this requires estimating $\mathbb{E}_{P(V|Z)}[\cdot]$ which may be hard if Z is continuous-valued or we do not have enough data for all possible values of Z. Assuming smoothness of this conditional loss, [5, Lemma 4.2] gives an upper bound for the worst-case loss with the empirical version as,

$$\widehat{\mathcal{R}}(\theta, \mathcal{U}_P^{\text{sub}}) = \inf_{\eta \geq 0, B \in \mathbf{R}_+^{n \times n}} \left\{ \frac{1}{\delta_0} \left(\frac{1}{n} \sum_{i=1}^{n} \left(\ell(\theta, V_i) - \frac{1}{n} \sum_{j=1}^{n} (B_{ij} - B_{ji}) - \eta \right)_+^2 \right)^{1/2} \right.$$
$$\left. + \frac{L}{\epsilon n^2} \sum_{i,j=1}^{n} \sum_{O \in Z} \|O_i - O_j\| B_{ij} + \eta, \text{ for any } \epsilon > 0 \right\}. \tag{7}$$

where η, B are dual variables. We will use this estimator for solving DRO with the uncertainty set $\mathcal{U}_P^{\text{sub}}$ as a particular instance of the proposed uncertainty set $\mathcal{U}_P^{\text{div}\cap\text{int}}$.

3.2 Our Approach for Robust Off-policy Evaluation in Contextual Bandits

Now we consider the problem of evaluating a policy from batch data. We have access to n tuples $\{V_i = (Z_i, T_i, Y_i)\}_i$ collected with a known stochastic policy that applies treatment T_i in context Z_i and observes the corresponding outcome Y_i. Further, assume that the tuples are i.i.d. as in the contextual bandit setup, where the joint distribution factorizes as $\prod_i P(Z_i)P(T_i|Z_i)P(Y_i|Z_i, T_i)$. Given data sampled from P, the goal in off-policy evaluation (OPE) is to evaluate the expected outcome $\mathbb{E}_Q[Y]$ under a distribution Q induced by following a new policy but in the same environment [6, 26]. Importantly, the difference between the two distributions is assumed to be only due to different policies i.e. $P(T|Z) \neq Q(T|Z)$ and the rest of the environment-related factors are the same. That is, in terms of the causal graph in Figure 1, only shift interventions on T are considered [27]. Departing from this assumption, our goal is to evaluate a policy under *new* environments characterized by unknown interventions on Z. Accounting for the uncertainty in the new marginal distribution of Z, the robust off-policy evaluation problem is to find the worst-case expected outcome,

$$\text{(Robust OPE)} \qquad \mathcal{R}(\mathcal{U}_P) = \sup_{Q \in \mathcal{U}_P} \mathbb{E}_{V \sim Q}[Y], \tag{8}$$

The uncertainty set $\mathcal{U}_P$ for the causal graph describing a general contextual bandit problem (Figure 1) is defined as

$$\mathcal{U}_P = \{Q \ll P \text{ s.t. } Q = \nu(Z)Q(T|Z)P(Y|Z,T), D[\nu(Z)\|P(Z)] \leq \delta\}. \tag{9}$$

where $Q(T|Z)$ is the policy to be evaluated and is considered to be known. Consider each distribution in the set $\mathcal{U}_P$, $Q = \nu(Z)Q(T|Z)P(Y|Z,T)$, which differs from the train distribution in the factors for Z and $T|Z$. To solve (8) for this Q, we first use the standard trick of importance sampling to account for the change in $T|Z$ due to the known policy $Q(T|Z)$. As a result, the set $\mathcal{U}_P$ now consists of shifts on Z alone, which is the same as the set $\mathcal{U}_P^{\text{div}\cap\text{int}}$ in (5). Thus, the problem reduces to solving $\sup_{Q \in \mathcal{U}_P^{\text{div}\cap\text{int}}} \mathbb{E}_{V \sim Q}[W \times Y]$ where W are the importance sampling weights, $W(T, Z) = Q(T|Z)/P(T|Z)$. We use the estimator (7) to solve this for subpopulation shifts in Z.

4 Experiments and Results

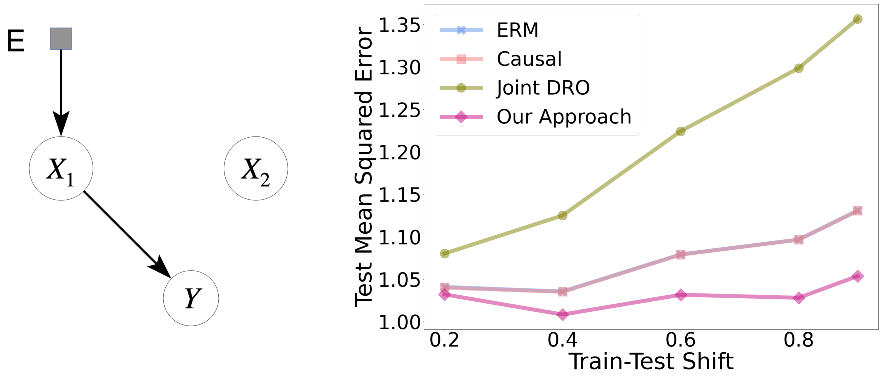
We report simulations for both the problems in Sections 3.1 and 3.2. The objective is to understand the achieved trade-off between utility and robustness compared to other robust learning approaches.

Supervised learning Consider the data generating process used in [5]. Structural equations are

$$Y = |X_1| + 1(X_1 \geq 0)\epsilon, \; \epsilon \sim \text{Normal}(0, 1),$$
$$X_1 \sim E \cdot \text{Uniform}[0, 1], X_2 \sim \text{Uniform}[-1, 1], E \sim \text{Bernoulli}(-1, 1, \delta).$$

There are two groups in the dataset, $X_1 \geq 0$ and $X_1 < 0$, with different outcome functions. Selection node E controls the relative proportion of each group. In the train data ($n = 2000$), the group with $X_1 < 0$ are in minority. The test data ($n = 2000$) has increasing proportion of minority group as δ is increased from 0.2 to 0.9. We compare with empirical risk minimization (ERM) $\mathbb{E}[Y|X_1, X_2]$, causal solution $\mathbb{E}[Y|pa(Y)] \equiv \mathbb{E}[Y|X_1]$, and Joint DRO [4] which considers bounded shifts on *all* variables Y, X_1, X_2. Since the shift is only on X_1, our approach fits models which are robust to bounded shifts in X_1 using the estimator (7). Joint DRO and our approach are trained with δ_0 in (6) equal to 0.2, which controls the desired robustness level. For all methods, we fit linear regression models with zero bias term. Note that, we choose a misspecified model to illustrate the effect of misspecification on the generalization of the causal solution as compared to other methods.

In Figure 2, we plot the mean squared error for each method on test sets with increasing minority group proportion. Note that the curves for ERM and Causal overlap as the uncorrelated variable X_2 does not change ERM solution much. ERM, as expected, is not robust and its test error increases with more shift. We observe that Joint DRO is overly conservative, i.e. has high error, as it aims to be robust to shifts in all variables. Although the causal solution should be robust to shifts on X_1, but due to model misspecification, the error is higher than the proposed approach and increases as minority group proportion increases. In summary, we observe that the proposed approach achieves low test error for both small and large shifts, whereas other approaches are either overly conservative (Joint DRO) or are not robust to model misspecification (Causal).



(a)

(b)

Figure 2: Supervised learning under marginal shift. (a) Causal graph for the data generating process. (b) Models are trained on a fixed train dataset. Mean squared error on test sets (y-axis) with varying levels of shift in X_1 (x-axis) is plotted. Required robustness level δ_0 in (6) is set to 0.2. We observe that solving DRO with shifts in X_1 alone results in a model that performs well across different shift magnitudes.

Robust OPE We consider the data generated according to the causal graph in Figure 1 with two features $Z := (X_1, X_2)$ in the context, binary treatment T and continuous outcome Y. Structural equations are:

$$X_1 \sim E \cdot \text{Normal}(10, 1), E \sim \text{Bernoulli}(-1, 1, \delta), X_2 \sim \text{Normal}(5, 1),$$
$$Y_{T=t} \sim \text{Normal}(X^\top w_t, 0.1^2), t \in \{0, 1\}, w_0 = [0.1, 0.1], w_1 = [0.1, 0.5]$$

We consider logistic policies, $T \sim \text{Bernoulli}(\sigma(X^\top \beta + \beta_0)), \beta = [0.1, 0.1], \sigma(x) = 1/(1 + e^{-x})$. The bias term in the train policy is changed from $\beta_0 = -1$ in train, whereas the policy to be evaluated in test environments has $\beta_0 = -0.5$. In addition, the marginal distribution of X_1 is changed, via change in E, by increasing δ in the test environment. We simulate $n = 2000$ samples in the train environment and use it for all methods. The objective is to estimate the robust value of the new policy. We compare with No Transfer that outputs the average value in train assuming there is no shift, inverse probability weighting (IPW) that corrects for shift in action distribution alone using importance sampling, Joint DRO that accounts for shifts in all variables Y, X_1, X_2, T. Note that Joint DRO for the case of contextual bandits has been proposed recently by [19] for sets defined using f-divergence.

In Figure 3, we plot the mean squared error between the estimated and the true policy value, evaluated using $n = 20000$ samples from the test environment. We observe that when the test environment is close to the train one, not accounting for the shift (No Transfer) performs well. But, as the shift increase, our approach does better. With large shifts, larger uncertainty sets are required. Thus, Joint DRO does better than the other methods. In summary, the proposed approach performs well when the shift is significant but not too large. This highlights the importance of choosing the desired robustness level appropriately which is a challenging problem for DRO methods in general.

5 Conclusion and Limitations

Distribution shifts have been represented in past work either using divergence metrics or using shift interventions on underlying causal models. Distributionally-robust optimization framework offers a unified way of learning models that are robust to such shifts. Here, we describe a way to represent shifts which are at the intersection of the two approaches. We argue that this gives a way to represent more realistic shifts and to find robust solutions which are less conservative. We show that under

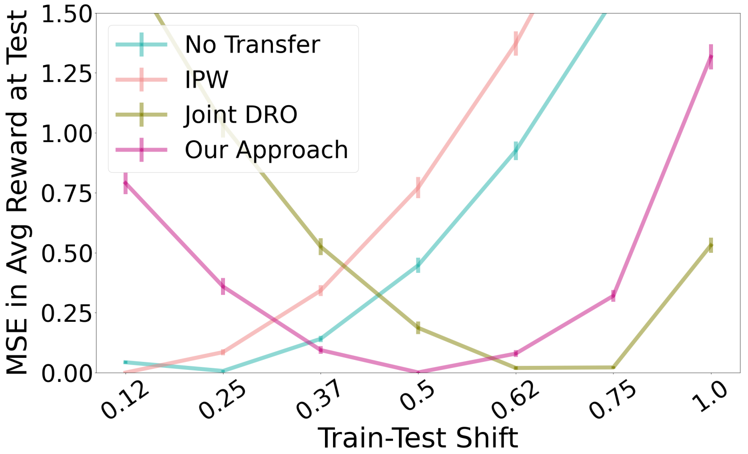


Figure 3: Off-policy evaluation in the contextual bandit setting in Figure 1. Robust value for the new policy is estimated only using data from the train environment. Mean squared error in the value estimate for test environments (y-axis) with varying levels of shift in Z (x-axis)is plotted. Robustness level in (6) is set to $\delta_0 = 0.8$. We observe that our approach performs well for moderate shifts. Mean and standard deviation for error bars are computed over 5 random intializations.

certain conditions the robust solution can be obtained efficiently for such shifts. For the supervised learning case, we provide a causal interpretation of an existing method [5]. We extend this to the case of off-policy evaluation to provide a novel estimator for the robust version of the problem.

One of the main limitations is the restrictive Assumptions 1(a) and 1(b). While knowledge of the causal graph allows precisely defining the uncertainty set, the graph may not be available in practice and has to be inferred using causal discovery methods along with inputs from domain experts. When there are unobserved confounders, further assumptions on the functional form of the causal relationships can be made to identify the worst-case loss. An interesting direction for further work is to develop techniques for solving DRO for shifts other than in marginal distributions (e.g. shift in the conditional distribution $Y|T, Z$ in Figure 1). Such development will enable expressing and learning robust models for a larger class of distribution shifts.

References

[1] M. Arjovsky, L. Bottou, I. Gulrajani, and D. Lopez-Paz. Invariant risk minimization. *arXiv preprint arXiv:1907.02893*, 2019.

[2] P. Bühlmann. Invariance, causality and robustness. *Statist. Sci.*, 35(3):404–426, 08 2020. doi: 10.1214/19-STS721. URL `https://doi.org/10.1214/19-STS721`.

[3] J. Correa and E. Bareinboim. A calculus for stochastic interventions: Causal effect identification and surrogate experiments. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, New York, NY, 2020. AAAI Press.

[4] J. Duchi and H. Namkoong. Learning models with uniform performance via distributionally robust optimization. *arXiv preprint arXiv:1810.08750*, 2018.

[5] J. C. Duchi, T. Hashimoto, and H. Namkoong. Distributionally robust losses against mixture covariate shifts. *Under review*, 2019.

[6] M. Dudík, D. Erhan, J. Langford, L. Li, et al. Doubly robust policy evaluation and optimization. *Statistical Science*, 29(4):485–511, 2014.

[7] P. M. Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1-2):115–166, 2018.

[8] R. J. Hernán MA. *Causal Inference: What If.* Boca Raton: Chapman & Hall/CRC, 2020. URL `https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/`.

[9] W. Hu, G. Niu, I. Sato, and M. Sugiyama. Does distributionally robust supervised learning give robust classifiers? In *International Conference on Machine Learning*, pages 2029–2037, 2018.

[10] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu. Towards deep learning models resistant to adversarial attacks. In *International Conference on Learning Representations*, 2018.

[11] S. Magliacane, T. van Ommen, T. Claassen, S. Bongers, P. Versteeg, and J. M. Mooij. Domain adaptation by using causal inference to predict invariant conditional distributions. In *Advances in Neural Information Processing Systems*, pages 10846–10856, 2018.

[12] J. Pearl. *Causality*. Cambridge university press, 2009.

[13] J. Pearl and E. Bareinboim. Transportability of causal and statistical relations: A formal approach. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*, 2011.

[14] J. Peters, P. Bühlmann, and N. Meinshausen. Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(5):947–1012, 2016.

[15] A. Raghunathan, S. M. Xie, F. Yang, J. Duchi, and P. Liang. Understanding and mitigating the tradeoff between robustness and accuracy. *arXiv*, pages arXiv–2002, 2020.

[16] M. Rojas-Carulla, B. Schölkopf, R. Turner, and J. Peters. Invariant models for causal transfer learning. *The Journal of Machine Learning Research*, 19(1):1309–1342, 2018.

[17] D. Rothenhäusler, N. Meinshausen, P. Bühlmann, and J. Peters. Anchor regression: heterogeneous data meets causality. *arXiv preprint arXiv:1801.06229*, 2018.

[18] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2014.

[19] N. Si, F. Zhang, Z. Zhou, and J. Blanchet. Distributional robust batch contextual bandits. *arXiv preprint arXiv:2006.05630*, 2020.

[20] A. Sinha, H. Namkoong, and J. Duchi. Certifying some distributional robustness with principled adversarial training. In *International Conference on Learning Representations*, 2018.

[21] M. Srivastava, T. Hashimoto, and P. Liang. Robustness to spurious correlations via human annotations. In *International Conference on Machine Learning*, 2020. URL `https://proceedings.icml.cc/static/paper_files/icml/2020/6535-Paper.pdf`.

[22] M. Staib and S. Jegelka. Distributionally robust deep learning as a generalization of adversarial training. In *NIPS workshop on Machine Learning and Computer Security*, 2017.

[23] M. Staib and S. Jegelka. Distributionally robust optimization and generalization in kernel methods. In *Advances in Neural Information Processing Systems*, pages 9134–9144, 2019.

[24] A. Subbaswamy, P. Schulam, and S. Saria. Preventing failures due to dataset shift: Learning predictive models that transport. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 3118–3127, 2019.

[25] R. Taori, A. Dave, V. Shankar, N. Carlini, B. Recht, and L. Schmidt. Measuring robustness to natural distribution shifts in image classification. *arXiv preprint arXiv:2007.00644*, 2020.

[26] P. Thomas and E. Brunskill. Data-efficient off-policy policy evaluation for reinforcement learning. In *International Conference on Machine Learning*, pages 2139–2148, 2016.

[27] J. Zhang and E. Bareinboim. Transfer learning in multi-armed bandit: a causal approach. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pages 1778–1780, 2017.