Notes on Conformal Risk Control

Catherine Chen cyc2152

December 2023

A. N. Angelopoulos, S. Bates, A. Fisch, L. Lei, and T. Schuster, “Conformal Risk Control.” arXiv, Apr. 29, 2023. Accessed: Dec. 14, 2023. [Online]. Available: http://arxiv.org/abs/2208.02814

Abstract: Control the expected value of any monotone loss function via conformal prediction, that is tight up to an $O(1/n)$ factor.

1 Introduction

1.1 Setting and Notation

Consider a *calibration dataset* $(X_i, Y_i)_{i=1,\ldots,n} \sim$ i.i.d. s.t. features vectors $X_i \in \mathcal{X}$ and response $Y_i \in \mathcal{Y}$. Conformal prediction seeks to bound the *miscoverage*, of a new test point (X_{n+1}, Y_{n+1})

$$\mathbb{P}\left(Y_{n+1} \notin \mathcal{C}\left(X_{n+1}\right)\right) \leq \alpha \tag{1}$$

where α is a user-specified error rate, and $\mathcal{C}$ is a function of the model and calibration data that produces the prediction set.

Objective: Provide guarantee called, *conformal risk control*, of the form

$$\mathbb{E}\left[\ell\left(\mathcal{C}\left(X_{n+1}\right), Y_{n+1}\right)\right] \leq \alpha, \tag{2}$$

for any bounded *loss function* ℓ that has an inverse relation with $\mathcal{C}(X_{n+1})$.

Remark: Recover conformal miscoverage guarantee with $\ell\left(\mathcal{C}\left(X_{n+1}\right), Y_{n+1}\right) = \mathbb{1}\left\{Y_{n+1} \notin \mathcal{C}\left(X_{n+1}\right)\right\}$.

Conformal risk control seeks to find a threshold $\hat{\lambda}$ that controls the proportion of missed classes:

$$\mathbb{E}\left[\ell\left(\mathcal{C}_{\hat{\lambda}}\left(X_{n+1}\right), Y_{n+1}\right)\right] = \mathbb{E}\left[1 - \frac{\left|Y_{n+1} \cap \mathcal{C}_{\lambda}\left(X_{n+1}\right)\right|}{\left|Y_{n+1}\right|}\right]$$

Note: the threshold λ will be defined s.t. it increases with $|\mathcal{C}_\lambda(x)|$; in other words, as λ grows, $\mathcal{C}_\lambda(x)$ becomes more conservative.

1.2 Algorithm and Preview of Main Results

Given base model f, post-process the predictions to produce a function $\mathcal{C}_\lambda(\cdot)$. The quality of the output of $\mathcal{C}_\lambda$ will be quantified by a loss function $\ell\left(\mathcal{C}_\lambda(x), y\right) \in (-\infty, B]$ for some $B < \infty$, that is a non-increasing as function of λ.

Goal: Choose $\hat{\lambda}$ based on the observed data $\{(X_i, Y_i)\}_{i=1}^n$ s.t. the risk control in (2) holds.

Consider an exchangeable collection of non-increasing, random functions $L_i : \Lambda \to (-\infty, B], i = 1, \ldots, n+1$. Assume $\lambda_{\max} \triangleq \sup \Lambda \in \Lambda$. Use the first n functions to choose a value of the parameter, $\hat{\lambda}$, s.t. the risk on the unseen function is controlled:

$$\mathbb{E}\left[L_{n+1}(\hat{\lambda})\right] \leq \alpha \tag{3}$$

Let $\widehat{R}_n(\lambda) = \frac{(L_1(\lambda)+\ldots+L_n(\lambda))}{n}$. Given any desired risk level upper bound $\alpha \in (-\infty, B)$, define

$$\hat{\lambda} = \inf\left\{\lambda : \frac{n}{n+1}\widehat{R}_n(\lambda) + \frac{B}{n+1} \leq \alpha\right\}. \tag{4}$$

When the set is empty, take $\hat{\lambda} = \lambda_{\max}$.

The proposed conformal risk control algorithm deploys $\hat{\lambda}$ on the new test point to achieve the guarantee in (3).

When the L_i are i.i.d. from a continuous distribution, the algorithm satisfies a tight lower bound illustrating that it is not too conservative,

$$\alpha - \frac{2B}{n+1} \leq \mathbb{E}\left[L_{n+1}(\hat{\lambda})\right] \leq \alpha.$$

2 Theory

2.1 Risk Control

Theorem 1. *Assume that $L_i(\lambda)$ is non-increasing in λ, right-continuous, and*

$$L_i\left(\lambda_{\max}\right) \leq \alpha, \quad \sup_\lambda L_i(\lambda) \leq B < \infty \text{ almost surely.} \tag{5}$$

Then

$$\mathbb{E}\left[L_{n+1}(\hat{\lambda})\right] \leq \alpha$$

Proof. Let $\widehat{R}_{n+1}(\lambda) = \frac{(L_1(\lambda)+\ldots+L_{n+1}(\lambda))}{n+1}$ and

$$\hat{\lambda}' = \inf\left\{\lambda \in \Lambda : \widehat{R}_{n+1}(\lambda) \leq \alpha\right\}.$$

Since $L_i(\lambda)$ is non-increasing in λ: $\inf_\lambda L_i(\lambda) = L_i\left(\lambda_{\max}\right) \leq \alpha$, thus $\hat{\lambda}'$ is well-defined almost surely.

By assumption $L_{n+1}(\lambda) \leq B$, we have $\widehat{R}_{n+1}(\lambda) = \frac{n}{n+1}\widehat{R}_n(\lambda) + \frac{L_{n+1}(\lambda)}{n+1} \leq \frac{n}{n+1}\widehat{R}_n(\lambda) + \frac{B}{n+1}$. Thus,

$$\left(\widehat{R}_{n+1}(\lambda) \leq\right) \underbrace{\frac{n}{n+1}\widehat{R}_n(\lambda) + \frac{B}{n+1} \leq \alpha}_{\inf\{LHS\} \implies \hat{\lambda}} \implies \underbrace{\widehat{R}_{n+1}(\lambda) \leq \alpha}_{\inf\{RHS\} \implies \hat{\lambda}'}.$$

Since $\hat{\lambda} = \inf\left\{\lambda : \frac{n}{n+1}\widehat{R}_n(\lambda) + \frac{B}{n+1} \leq \alpha\right\}$, taking the infimum on both sides implies $\hat{\lambda}' \leq \hat{\lambda}$ when the LHS holds for some $\lambda \in \Lambda$.

When the LHS is above α for all $\lambda \in \Lambda$, the set, $\left\{\lambda : \frac{n}{n+1}\widehat{R}_n(\lambda) + \frac{B}{n+1} \leq \alpha\right\}$ is empty, thus by definition, we take $\hat{\lambda} = \lambda_{\max} \implies \hat{\lambda} \geq \hat{\lambda}'$. Thus, $\hat{\lambda}' \leq \hat{\lambda}$ almost surely. Since $L_i(\lambda)$ is non-increasing in λ,

$$\mathbb{E}\left[L_{n+1}(\hat{\lambda})\right] \leq \mathbb{E}\left[L_{n+1}\left(\hat{\lambda}'\right)\right] \tag{6}$$

Let E be the multiset of loss functions $\{L_1, \ldots, L_{n+1}\}$. Then $\hat{\lambda}'$ is a function of $E \iff \hat{\lambda}'$ is a constant conditional on E. Additionally, $L_{n+1}(\lambda) \mid E \sim \text{Uniform}\left(\{L_1, \ldots, L_{n+1}\}\right)$ by exchangeability. Combined with the right-continuity of L_i, it can be shown that

$$\mathbb{E}\left[L_{n+1}\left(\hat{\lambda}'\right) \mid E\right] = \frac{1}{n+1}\sum_{i=1}^{n+1} L_i\left(\hat{\lambda}'\right) \leq \alpha.$$

By the law of total expectation and (6), the proof is complete. □

2.2 A tight risk lower bound

2.3 Conform prediction reduces to risk control

2.4 Controlling general loss function