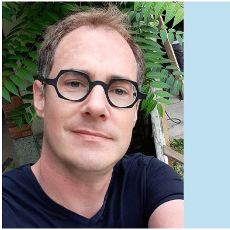
Improving fairness in online learning

Christophe Giraud

Laboratoire de Mathématiques d'Orsay,
Université Paris Saclay

Heidelberg 2023

Collaborators and references



A unified approach to fair online learning via Blackwell approachability

E. Chzhen, C. Giraud, G. Stoltz; NeurIPS 2021 (spotlight).

Small Total-Cost Constraints in CBwK, with Application to Fairness

E. Chzhen, C. Giraud, Z. Li, G. Stoltz; arXiv:2305.15807

Parameter-free projected gradient descent

E. Chzhen, C. Giraud, G. Stoltz; arXiv:2305.19605

Collaborators and references

The price of unfairness in linear bandits with biased feedback

S. Gaucher, A. Carpentier, C. Giraud; NeurIPS 2022.

Fairness in Machine Learning: a major societal concern

Machine Learning is ubiquitous in daily life, and it is used for sensitive decisions such as

- admission in university,
- bank loan,
- job recruitment,
- justice decision,
- ...

Promises of ML in decision-making

Promises of ML in decision-making

ML can be more objective and more fair than humans, as algorithms can

- incorporate more data, and more factors in a complex analysis,
- and are not subject to personal biases, tiredness, emotional factors, etc

Hungry judge effect: some historical court data for grant to parole decisions (release from jail) have revealed that, irrespective of the gravity of the crimes

- $\approx 65\%$ of positive decisions in the early morning or afternoon,
- close to 0% of positive decisions before lunch.

Actuality of ML decision-making

Discriminations also happen in ML prediction

Many ML systems have been shown to produce unfair outcomes.

Some famous past examples:

- **Hiring AI** from Amazon was discriminating against female candidate on some jobs
- **Google Ad** was proposing higher-paying executive jobs more likely to men than women
- **COMPAS** was falsely predicting recidivism twice more likely for African-American than for Caucasian-American.

Where does the unfairness come from?

Main potential causes of unfairness in data science

- [intentional discrimination]
- **historical biases in learning datasets**
- inadvertent bias in evaluations (biased proxy)
- inadvertent bias from data sampling: learning dataset not representative of the target population
- **inadvertent bias from algorithm objectives: focus on the benefit for majority group**

How can we mitigate these issues?

This talk: some possible directions for improving fairness in online learning

1. Causal Fairness
2. Statistical Fairness
   - Adversarial setting
   - Stochastic setting

Contextual online setting

Covariate and sensitive attribute

Each request is characterized by a covariate $x \in \mathcal{X}$ (observed) and a sensitive attribute $s \in \{-1, +1\}$ (observed or not).

Informal description of a typical setting

A each epoch $t = 1, 2, \ldots$

- The Learner observes a context (x_t, s_t) or x_t only
- The Learner performs an action (or prediction) a_t
- The Learner observes a feedback y_t and suffers a regret r_t (stochastic or adversarial)

Goal of the learner

To minimize the cumulative regret $\sum_t r_t$, while complying to some fairness criteria (and possibly some other constraints).

1- A causal point of view

Causal fairness: identifying causes of unfairness

Causal fairness: general principle

- The relations between attributes (X, S) and their influence on feedback Y is modeled by structural equations
- The objective is to remove all discriminatory influences of sensitive attribute on the action/prediction

Example: No unresolved discrimination

Design an action/prediction such that no path from the sensitive attribute to the action exists, except via non-discriminatory variables (resolving variables).

Caveat: the notions of causal fairness heavily rely on the causal model. The accuracy of this model is critical, and learning it can be problematic.

A very (not so) simple case

The price of unfairness in linear bandits with biased feedback

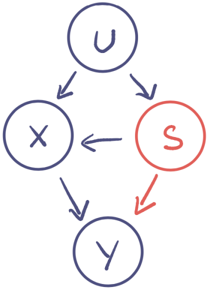
S. Gaucher, A. Carpentier, C. Giraud; NeurIPS 2022.

Biased linear feedbacks

- biased feedback
  $y_t = x_t^\top \beta^* + s_t b^* + \xi_t$
- unobserved regret
  $r_t = \max_{x \in \mathcal{X}} x^\top \beta^* - x_t^\top \beta^*$,

with β^* and b^* unknown.

Resolving variables: all x_t



Price for learning the bias

- **Worst case:** the minimax regret scales as $T^{2/3}$ instead of $\sqrt{T}$ for the unbiased case
- **Asymptotic regret:** scales as $\Delta^{-2} \log(T)$ instead of $\Delta^{-1} \log(T)$ in some cases.

2- Statistical fairness

Statistical fairness

General principle

To comply to some fairness criteria at the sub-population level (statistical notions)

Example 1: equalized Odds

$$A \perp\!\!\!\perp S \mid Y$$

- Equalized Odds encodes a notion of meritocracy
- There are many variants

Caveat: strongly subject to biases in learning datasets

Statistical fairness: Demographic parity

Example 2: Demographic parity

$$A \perp\!\!\!\perp S$$

Demographic parity promotes *diversity* and can be related to affirmative action policies.

Caveat: the feedback Y is not taken into account

Finding a balance between different notions

Relaxing fairness criteria

- Fairness criteria are imperfect mathematical transposition of qualitative ideas;
- Evaluations of fairness criteria are subjected to uncertainties;
- Some fairness criteria are incompatible;

so, it is wise to

- introduce some quantitive measures of violation of the fairness criteria;
- seek for a good trade-off between different fairness criteria and regret (Pareto frontier).

An instantiation in online learning

Fairness cost

Fairness criteria can be encoded as vector valued cost constraints.

Example: demographic parity

The empirical demographic parity criteria (for $a_t \in \{0,1\}$)

$$\left| \frac{1}{p_1 T} \sum_{t \leq T; s_t = 1} a_t - \frac{1}{p_{-1} T} \sum_{t \leq T; s_t = -1} a_t \right| = \tilde{O}(T^{-1/2})$$

can be encoded as

$$\sum_{t \leq T} c_t = \tilde{O}(\sqrt{T}) \quad \text{with} \quad c_t := \begin{bmatrix} a_t s_t / p_{s_t} \\ -a_t s_t / p_{s_t} \end{bmatrix}.$$

Our contributions

Informal objective

$$\min_{\sum_{t\leq T} c_t \leq \tilde{O}(\sqrt{T})} \sum_{t\leq T} r_t .$$

Two points of view

1. **In adversarial setting:**
   - the fair learning problem can be formulated as a contextual approachability problem,
   - Blackwell theory can be adapted to handle this setting.

2. **In stochastic bandit setting:**
   - the fairness objective falls into the Contextual Bandit with Knapsack (CBwK) framework,
   - the theory for CBwK must be improved to handle $\tilde{O}(\sqrt{T})$ constraints (and signed cost).

Adversarial Setting :

a Contextual Blackwell Approachability Perspective

A unified approach to fair online learning via Blackwell approachability.

E. Chzhen, C. Giraud, G. Stoltz; NeurIPS 2021 (spotlight).

Online learning setting: formal description

We model our fair online learning problem as a contextual learning game between the Learner and Nature.

Stochastic attributes (context)

At each time t, the attributes (x_t, s_t) are sampled according to $\mathbf{Q}$, independently from the past.

Nature (un)awareness

Let G denotes Nature (un)awareness mapping

- Nature *awareness* $G(x, s) = (x, s)$,
- Nature *unawareness*: $G(x, s) = x$.

Nature is an adverse player

At each time t, Nature observes $G(x_t, s_t)$ and outputs an adversarial feedback y_t.

Fair online learning as a contextual approachability problem

Encoding the objectives of the learner

We can encode the learning objectives (vanishing-regret, demographic parity, etc) via

- a vector-valued payoff function $\mathbf{m}(a_t, y_t, x_t, s_t)$
- and a target set $\mathcal{C}$.

The learning objective is to comply to $\frac{1}{T}\sum_{t=1}^{T}\mathbf{m}(a_t, y_t, x_t, s_t) \longrightarrow \mathcal{C}$.

Examples of targets (to be combined)

Criterion	Vector payoff function $\mathbf{m}$	Closed convex target set $\mathcal{C}$
Demographic parity	$\mathbf{m}_{\mathrm{DP}}(a,s) = \left(\frac{a}{p_{-1}}\,\mathbf{1}_{s=-1},\ \frac{a}{p_1}\,\mathbf{1}_{s=1}\right)$	$\mathcal{C}_{\mathrm{DP}} = \{(u,v)\in\mathbb{R}^2:\ \lvert u-v\rvert\le\delta\}$
No-regret	$\mathbf{m}_{\mathrm{reg}}(a,y,x,s) = \left(f(a,y,x,s) - f(a',y,x,s)\right)_{a'\in\mathcal{A}}$	$\mathcal{C}_{\mathrm{reg}} = [0,+\infty)^N$
Group-calibration	$\mathbf{m}_{\mathrm{gr\text{-}cal}}(a,y,s) = \left((a'-y)\,\mathbf{1}_{s=s'}/\gamma_{s'}\right)_{a'\in\mathcal{A},\ s'\in\mathcal{S}}$	$\mathcal{C}_{\mathrm{gr\text{-}cal}} = \{\mathbf{v}\in\mathbb{R}^{N\lvert\mathcal{S}\rvert}:\ \lVert\mathbf{v}\rVert_1\le\varepsilon\}$
Equalized payoffs	$\mathbf{m}_{\mathrm{eq\text{-}pay}}(a,y,x,s) = \left(\frac{f(a,y,x,s')}{\gamma_{s'}}\,\mathbf{1}_{s=s'}\right)_{s'\in\mathcal{S}}$	$\mathcal{C}_{\mathrm{eq\text{-}pay}} = \{(u,v)\in\mathbb{R}^2:\ \lvert u-v\rvert\le\varepsilon\}$

Online learning setting

Learning setting

For $t = 1, 2, \ldots$

1. Simultaneously,
   - the Learner chooses $(\mathbf{p}_t^x)_{x \in \mathcal{X}}$ based on $(\mathbf{m}_\tau, x_\tau, s_\tau)_{\tau \le t-1}$
   - Nature chooses $\left(\mathbf{q}_t^{G(x,s)}\right)_{(x,s) \in \mathcal{X} \times \mathcal{S}}$ based on $(a_\tau, y_\tau, x_\tau, s_\tau)_{\tau \le t-1}$
2. (x_t, s_t) are sampled according to $\mathbf{Q}$, independently from the past
3. Simultaneously
   - the Learner observes x_t, and picks an action $a_t \in \mathcal{A}$ according to $\mathbf{p}_t^{x_t}$
   - Nature observes $G(x_t, s_t)$, and picks $y_t \in \mathcal{Y}$ according to $\mathbf{q}_t^{G(x_t, s_t)}$
4. The Learner observes the payoff $\mathbf{m}_t = \mathbf{m}(a_t, y_t, x_t, s_t)$ and (x_t, s_t), while Nature observes (a_t, y_t, x_t, s_t).

Aim: The Learner wants to ensure that $\bar{\mathbf{m}}_T := \frac{1}{T} \sum_{t=1}^{T} \mathbf{m}_t \to \mathcal{C}$ a.s. for some target set $\mathcal{C}$.

Online learning setting

Learning setting

For $t = 1, 2, \ldots$

1. Simultaneously,
   - the Learner chooses $(\mathbf{p}_t^x)_{x \in \mathcal{X}}$ based on $(\mathbf{m}_\tau, x_\tau, s_\tau)_{\tau \leq t-1}$
   - Nature chooses $\left(\mathbf{q}_t^{G(x,s)}\right)_{(x,s) \in \mathcal{X} \times \mathcal{S}}$ based on $(a_\tau, y_\tau, x_\tau, s_\tau)_{\tau \leq t-1}$
2. (x_t, s_t) are sampled according to $\mathbf{Q}$, independently from the past
3. Simultaneously
   - the Learner observes x_t, and picks an action $a_t \in \mathcal{A}$ according to $\mathbf{p}_t^{x_t}$
   - Nature observes $G(x_t, s_t)$, and picks $y_t \in \mathcal{Y}$ according to $\mathbf{q}_t^{G(x_t, s_t)}$
4. The Learner observes the payoff $\mathbf{m}_t = \mathbf{m}(a_t, y_t, x_t, s_t)$ and (x_t, s_t), while Nature observes (a_t, y_t, x_t, s_t).

Aim: The Learner wants to ensure that $\bar{\mathbf{m}}_T := \dfrac{1}{T}\sum_{t=1}^{T} \mathbf{m}_t \to \mathcal{C}$ a.s. for some target set $\mathcal{C}$.

Online learning setting

Learning setting

For $t = 1, 2, \ldots$

1. Simultaneously,
   - the Learner chooses $(\mathbf{p}_t^x)_{x\in\mathcal{X}}$ based on $(\mathbf{m}_\tau, x_\tau, s_\tau)_{\tau\leq t-1}$
   - Nature chooses $\left(\mathbf{q}_t^{G(x,s)}\right)_{(x,s)\in\mathcal{X}\times\mathcal{S}}$ based on $(a_\tau, y_\tau, x_\tau, s_\tau)_{\tau\leq t-1}$
2. (x_t, s_t) are sampled according to $\mathbf{Q}$, independently from the past
3. Simultaneously
   - the Learner observes x_t, and picks an action $a_t \in \mathcal{A}$ according to $\mathbf{p}_t^{x_t}$
   - Nature observes $G(x_t, s_t)$, and picks $y_t \in \mathcal{Y}$ according to $\mathbf{q}_t^{G(x_t,s_t)}$
4. The Learner observes the payoff $\mathbf{m}_t = \mathbf{m}(a_t, y_t, x_t, s_t)$ and (x_t, s_t), while Nature observes (a_t, y_t, x_t, s_t).

Aim: The Learner wants to ensure that $\bar{\mathbf{m}}_T := \dfrac{1}{T}\displaystyle\sum_{t=1}^{T} \mathbf{m}_t \to \mathcal{C}$ a.s. for some target set $\mathcal{C}$.

Assumption: fast enough sequential estimation of $\mathbf{Q}$

The Player can build estimators $(\hat{\mathbf{Q}}_t)_{t\geq 1}$ of the unknown distribution $\mathbf{Q}$ such that

$$\mathbb{E}\left[\mathrm{TV}^2(\hat{\mathbf{Q}}_t,\mathbf{Q})\right] \leq c\ \big(\log(t)\big)^{-3} \qquad \forall t \geq 2 \tag{1}$$

Theorem : Contextual Blackwell approachability

If $\mathcal{C} \subset \mathbb{R}^d$ is closed convex, $\mathbf{m}$ is bounded, and (1) is satisfied, then

$\exists(\mathbf{p}_t^x)_{x\in\mathcal{X},t\geq 1}$ such that $\forall(\mathbf{q}_t^{G(x,s)})_{(x,s)\in\mathcal{X}\times\{0,1\},t\geq 1}$ we have $\bar{\mathbf{m}}_T \overset{a.s.}{\to} \mathcal{C}$

if and only if $\forall(\mathbf{q}^{G(x,s)})_{(x,s)\in\mathcal{X}\times\{0,1\}}$ $\exists(\mathbf{p}^x)_{x\in\mathcal{X}}$ such that

$$\mathbf{m}(\mathbf{p},\mathbf{q},\mathbf{Q}) := \int_{\mathcal{X}\times\mathcal{S}} \mathbf{m}\big(\mathbf{p}^x,\mathbf{q}^{G(x,s)},x,s\big)\,d\mathbf{Q}(x,s) \in \mathcal{C}$$

Contextual Blackwell strategy

Set $\mathbf{m}(\mathbf{p},\mathbf{q},\hat{\mathbf{Q}}_t) := \int \mathbf{m}\big(\mathbf{p}^x,\mathbf{q}^{G(x,s)},x,s\big)\,d\hat{\mathbf{Q}}_t(x,s)$. At stage $t+1$, choose

$$(\mathbf{p}_{t+1}^x)_{x\in\mathcal{X}} \in \underset{(\mathbf{p}^x)_x}{\mathrm{argmin}}\ \max_{(\mathbf{q}^{G(x,s)})_{x,s}} \langle \bar{\mathbf{m}}_t - \Pi_{\mathcal{C}}\bar{\mathbf{m}}_t, \mathbf{m}(\mathbf{p},\mathbf{q},\hat{\mathbf{Q}}_t)\rangle$$

Caveats

Caveat 1: the target set $\mathcal{C}$ has to be known

The results can be extended (at the price of some technicalities) to the case where we only have a consistent super-estimate $\hat{\mathcal{C}}_t$ of $\mathcal{C}$.

Caveat 2: computational cost of projection

Computing the projection $\Pi_{\mathcal{C}}$ can be computationally expensive.

Caveat 3: pessimistic Pareto frontier and slow rates

- The adversarial setting leads to pessimistic Pareto frontier (trade-off) between the different criteria;
- The rates are governed by the estimation rate $\text{TV}(\hat{\mathbf{Q}}_t, \mathbf{Q})$, which is typically slow outside the finite case.

Stochastic setting:

A Contextual Bandit with Knapsack perspective

Small Total-Cost Constraints in CBwK, with Application to Fairness

E. Chzhen, C. Giraud, Z. Li, G. Stoltz; arXiv:2305.15807

Stochastic setting

Learning problem

- The learner observes $\tilde{x}_t = (x_t, s_t) \overset{\text{i.i.d.}}{\sim} \mathbf{Q}$
- The learner chooses a policy $\pi_t : \tilde{\mathcal{X}} \to \mathcal{P}(\mathcal{A})$, and picks an action $a_t \sim \pi_t(\tilde{x}_t)$,
- The learner receives a feedback y_t and a fairness cost c_t such that

$$\mathbb{E}[y_t | \mathcal{F}_t] = f(\tilde{x}_t, a_t) \quad \text{and} \quad \mathbb{E}[c_t | \mathcal{F}_t] = c(\tilde{x}_t, a_t).$$

- The learner suffers a regret $r_t = \mathrm{OPT} - y_t$ (described below).

Example: Demographic Parity

$$c(\tilde{x}_t, a_t) = \begin{bmatrix} a_t s_t / p_{s_t} \\ -a_t s_t / p_{s_t} \end{bmatrix}.$$

Optimal policy and regret

Optimal static policy and regret

The optimal static feedback is

$$\mathrm{OPT}(\mathbf{Q}, f, c) := \max_{\pi \,:\, \mathbb{E}_{\mathbf{Q}}\left[\sum_{a\in\mathcal{A}} c(\tilde{X},a)\pi_a(\tilde{X})\right]\leq\delta_T} \mathbb{E}_{\mathbf{Q}}\left[\sum_{a\in\mathcal{A}} f(\tilde{X}, a)\pi_a(\tilde{X})\right]$$

and the regret is

$$r_t = \mathrm{OPT}(\mathbf{Q}, f, c) - y_t.$$

Learning Objective

Minimize the cumulative regret $\sum_{t\leq T} r_t$ while complying to the fairness constraint $\sum_{t\leq T} c_t \leq T\delta_T$.

Fairness as CBwK

CBwK problem

1. we recognize a Contextual Bandit with Knapsack (CBwK) problem
2. but state of the art theory can only handle $\delta_T = T^{-1/4}$ (or $\mathcal{X}$ finite), which is too large for fairness constraints, where we typically wish to have $\delta_T = \tilde{O}(T^{-1/2})$

Learning assumption

Assumption: UCB and LCB

We can built UCB and LCB such that with probability $\geq 1-\delta$

$$\hat{f}_t^{\text{UCB}}(.,.) \approx f(.,.) + \tilde{O}_\delta(1/\sqrt{t})$$
$$\hat{c}_t^{\text{LCB}}(.,.) \approx c(.,.) + \tilde{O}_\delta(1/\sqrt{t})$$

Examples

Linear or logistic model : when

$$f(x,a) = \eta(\varphi(x,a)^T\theta_a) \quad \text{and} \quad c(x,a) = \eta(\psi(x,a)^T\beta_a),$$

with $\eta(u) = u$ or $\eta(u) = e^u/(1+e^u)$, we can use variant of LinUCB or LogisticUCB1.

A first idea

Idea1: playing empirical optimal static policy

Choose a_t according to a policy $\hat{\pi}_t$ maximizing $\mathrm{OPT}(\hat{\mathbf{Q}}_t, \hat{f}_t^{\mathrm{UCB}}, \hat{c}_t^{\mathrm{LCB}})$.

Issues

1. The analysis of
$$\mathrm{OPT}(\mathbf{Q}, f, c) - \mathrm{OPT}(\hat{\mathbf{Q}}_t, \hat{f}_t^{\mathrm{UCB}}, \hat{c}_t^{\mathrm{LCB}})$$
produces some $\mathrm{TV}(\hat{\mathbf{Q}}_t, \mathbf{Q})$ terms, leading to slow rates / large fairness violation.
2. Solving $\mathrm{OPT}(\hat{\mathbf{Q}}_t, \hat{f}_t^{\mathrm{UCB}}, \hat{c}_t^{\mathrm{LCB}})$ is computationally expensive

Lagrangian version

Lagrangian formulation

$$\begin{aligned}\mathrm{OPT}(\mathbf{Q}, f, c) &= \max_{\pi\ :\ \mathbb{E}_{\mathbf{Q}}\left[\sum_{a\in\mathcal{A}} c(\tilde{X},a)\pi_a(\tilde{X}))\right]\leq\delta_T} \mathbb{E}_{\mathbf{Q}}\left[\sum_{a\in\mathcal{A}} f(\tilde{X}, a)\pi_a(\tilde{X})\right] \\ &= \min_{\lambda\geq 0}\max_{\pi} \mathbb{E}_{\mathbf{Q}}\left[\sum_{a\in\mathcal{A}} \pi_a(\tilde{X})\left(f(\tilde{X}, a) - \langle \lambda, c(\tilde{X}, a) - \delta_T\rangle\right)\right] \\ &= \min_{\lambda\geq 0} \ \mathbb{E}_{\mathbf{Q}}\left[\max_{a\in\mathcal{A}}\left\{f(\tilde{X}, a) - \langle \lambda, c(\tilde{X}, a) - \delta_T\rangle\right\}\right]\end{aligned}$$

Two immediate benefits

1. for a fixed λ the problem is separable, and $\mathbf{Q}$ can be forgotten;
2. we only need to learn the optimal $\lambda^* \in \mathbb{R}^d \implies$ parametric rates. ☺

High-level algorithm: Primal-dual descent-ascent

Iterate

- **full optimisation on primal variable:** pick

$$a_t \in \operatorname*{argmax}_{a \in \mathcal{A}} \left\{ \hat{f}_t(\tilde{X}, a) - \langle \lambda_{t-1}, \hat{c}_t(\tilde{X}, a) - \delta_T \rangle \right\}$$

- **projected subgradient step on dual variable:** update

$$\lambda_t = \left(\lambda_{t-1} + \gamma \left(\hat{c}_t(\tilde{X}, a) - \delta_T \right) \right)_+$$

Issues

1. **Benign issue:** we must replace δ_T by $\delta'_T = \delta_T - \tilde{O}(1/\sqrt{T})$ to prevent from violation of the fairness criteria due to random fluctuations
2. **Major issue:** to satisfy the constraints, we need to set $\gamma \approx \|\lambda^*\|/\sqrt{T}$ ☹

Adaptative algorithm

Adaptive version

Iterate: for $t \geq 1$

- Pick $a_t \in \text{argmax}_{a \in \mathcal{A}} \left\{ \hat{f}_t(\tilde{X}, a) - \langle \lambda_{t-1}, \hat{c}_t(\tilde{X}, a) - \delta'_T \rangle \right\}$
- Update $\lambda_t = \left(\lambda_{t-1} + \frac{2^k}{\sqrt{T}} \left(\hat{c}_t(\tilde{X}, a) - \delta'_T \right) \right)_+$

Until $\left\| \left(\sum_{\tau = T_k}^{t} c_\tau - (t - T_k + 1)\delta'_T \right)_+ \right\| > \tilde{O}(\sqrt{T})$

Then: increase k by one, set $T_k = t + 1$, and **iterate** again.

Theory

Regret bound

For $\delta_T \geq \tilde{O}(T^{-1/2})$, the above algorithm fulfills with probability at least $1-\delta$

$$\sum_{t \leq T} r_t \leq \tilde{O}_\delta\left((1+\|\lambda^*\|)\sqrt{T}\right) \qquad \sum_{t \leq T} c_t \leq \delta_T T.$$

Suitable for fairness constraints ☺

Optimality?

A proof scheme suggests that this regret is optimal.

Concluding remarks

- Fairness in decision-making is an important topic;
- The statistical community has an important role to play for providing
  - conceptual ideas
  - competitive algorithms with provable performances
  - theoretical insights
  - **education of the next generation of data scientists**