Stochastic Online Linear Regression: the Forward Algorithm to Replace Ridge

Reda Ouhamma[1], Odalric Maillard[1], Vianney Perchet[2]

[1]Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France
[2]Crest, Ensae & Criteo AI Lab

Accepted at NeurIPS 2021

Séminaire MIA

October, 2022

Talk roadmap

- Online linear regression, adversarial setting
- Existing analysis and limitations
- Stochastic setting, new analysis
- Application for linear bandits
- Experiments
- Unregularized forward algorithm
- Conclusion

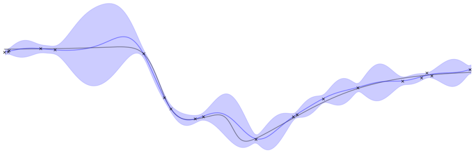
Online linear regression

Adversarial setting

Nature: Provides $(x_t)_{t\geq 1} \in \mathbb{R}^d$ and $y_t \in [-Y, Y]$ for a fixed $Y \in \mathbb{R}_+$

Interaction protocol: At step t

- Nature provides $x_t \in \mathbb{R}^d$
- The learner chooses $\hat{\theta}_{t-1}$ and predicts $\hat{y}_t = x_t^\top \hat{\theta}_{t-1}$
- Nature chooses y_t and shows it to the learner
- The learner suffers loss $\ell_t = (y_t - \hat{y}_t)^2$



Objective: minimize the cumulative regret

$$R_T^{\mathcal{A}} = L_T^{\mathcal{A}} - \min_\theta L_T(\theta) = \sum_{t=1}^T \ell_t^{\mathcal{A}} - \min_\theta L_T(\theta)$$

Literature

1/2

Online ridge regression

Input: initial parameter $\theta^r \in \mathbb{R}^d$, and regularization $\lambda \in \mathbb{R}_+$
for $t = 1, 2, \ldots, T$
  Observe input x_t
  predict $\hat{y}_t = x_t^T \theta_{t-1}^r \in \mathbb{R}$
  observe y_t
  incur $\ell_t \in \mathbb{R}$
  $\theta_t^r \in \operatorname{argmin}_\theta L_t(\theta) + \lambda ||\theta||_2^2$
end for

The optimization has a closed form solution:

$$\theta_t^r = \underbrace{\left(\lambda I + \sum_{q=1}^{t} x_q x_q^\top\right)^{-1}}_{\stackrel{\text{def}}{=} G_t(\lambda)^{-1}} \underbrace{\sum_{q=1}^{t} x_q y_q}_{\stackrel{\text{def}}{=} b_t}.$$

Literature

2/2

The forward algorithm (aka the Vovk-Azoury-Warmuth forecaster) [Vovk, 2001, Azoury and Warmuth, 2001]

Input: initial $\theta_0 \in \mathbb{R}^d$, and regularization $\lambda \in \mathbb{R}_+$
for $t = 1, 2, \ldots, T$
  Observe input x_t
  $\theta^f_{t-1} \in \text{argmin}_\theta\, L_{t-1}(\theta) + \lambda||\theta||_2^2$ $+(x_t^T\theta)^2$
  predict $\hat{y}_t = x_t^T \theta^f_{t-1} \in \mathbb{R}$
  observe y_t
  incur $\ell_t \in \mathbb{R}$
end for

The optimization has a closed form solution:

$$\theta^f_t = G^{-1}_{t+1} b_t = (\underbrace{G_t}_{\text{Same as Ridge}} + x_{t+1}x_{t+1}^\top)^{-1} \underbrace{b_t}_{\text{Same as Ridge}}$$

Regret bounds

Existing results

Online ridge regression achieves the bound:

$$L_T^r - \min_\theta \left(L_T(\theta) + \lambda \|\theta\|_2^2 \right) \leq 4 \left(Y^r \right)^2 d \log \left(1 + \frac{TX^2}{\lambda d} \right),$$

where $X = \max_{1 \leq t \leq T} \|x_t\|_2$, and $Y^r = \max_{1 \leq t \leq T} \left\{ |y_t|, \left| \boldsymbol{x}_t^\top \boldsymbol{\theta}_{t-1} \right| \right\}$.

Forward regression (see [Vovk, 2001]) achieves:

$$L_T^f - \min_\theta \left(L_T(\theta) + \lambda \|\theta\|_2^2 \right) \leq \left(Y^f \right)^2 d \log \left(1 + \frac{TX^2}{\lambda d} \right),$$

where $X = \max_{1 \leq t \leq T} \|x_t\|_2$, and $Y^f = \max_{1 \leq t \leq T} |y_t|$.

Limitations of the analysis

For Ridge regression $\quad R_T^r \leq 4\left(Y^r\right)^2 d \log\left(1 + \frac{TX^2}{\lambda d}\right) + \frac{\lambda (Y^r)^2 T}{\lambda_{r_T}(G_T(0))}$

Two issues:

1. Bound on labels: $\quad$ we have $Y^r = \max_{1 \leq t \leq T} \left\{|y_t|, \left|\boldsymbol{x}_t^\top \boldsymbol{\theta}_{t-1}\right|\right\}$

- The predictions $\boldsymbol{x}_t^\top \boldsymbol{\theta}_{t-1}$ are not necessarily bounded
- The latter cannot be mitigated unless Y is known

2. Regularization: $\quad \lambda_{r_T}$ is the smallest eigenvalue of $\lambda I_d + \sum_{q=1}^T x_q x_q^\top$.

- To cancel out $\frac{\lambda (Y^r)^2 T}{\lambda_{r_T}(G_T(0))}$, λ should be of order $1/T$
- $\lambda \sim 1/T$ yields $\quad R_T^r \lesssim 8(Y^r)^2 \log(T) \quad$ and $\quad R_T^f \lesssim 2Y^2 \log(T)$
- [Gaillard et al., 2019] Provides a tight lower bound of $dY^2 \log(T)$

Stochastic online linear regression:

Nature: Provides $(x_t)_{t\geq 1} \in \mathbb{R}^d$ and $\theta_* \in \mathbb{R}^d$

Data generating process: $y_t \stackrel{\text{def}}{=} x_t^\top \theta_* + \epsilon_t \in \mathbb{R}$ where the noise sequence is σ-sub-Gaussian, *i.e.*

$$\forall t \geq 1, \forall \gamma \in \mathbb{R} : \quad \mathbb{E}\left[\exp(\gamma\epsilon)\right] \leq \exp(\sigma^2\gamma^2/2).$$

Interaction protocol: At step t

- Nature provides $x_t \in \mathbb{R}^d$
- The learner chooses $\hat{\theta}_{t-1}$ and predicts $\hat{y}_t = x_t^\top \hat{\theta}_{t-1}$
- Nature chooses y_t and shows it to the learner
- The learner suffers loss $\ell_t = (y_t - \hat{y}_t)^2$

Objective: minimize the cumulative regret

$$R_T^{\mathcal{A}} = L_T^{\mathcal{A}} - \min_\theta L_T(\theta) = \sum_{t=1}^T \ell_t^{\mathcal{A}} - \min_\theta L_T(\theta)$$

Limitations of the analysis

1/2: Too generic

In the final step of the analysis we have

For ridge regression:

$$L_T^{\mathrm{r}} - \min_\theta \left(L_T(\theta) + \lambda \|\theta\|_2^2 \right) \leq \sum_{t=1}^{T} \underbrace{\left(x_t^\top \theta_{t-1} - y_t \right)^2 x_t^\top G_t^{-1} x_t}_{\text{first term}}$$

$$\leq 4 (Y^r)^2 \sum_{t=1}^{T} x_t^\top G_t^{-1} x_t$$

Limitations of the analysis

1/2: Too generic

In the final step of the analysis we have

For ridge regression:

$$L_T^{\mathrm{r}} - \min_\theta \left(L_T(\theta) + \lambda \|\theta\|_2^2 \right) \le \sum_{t=1}^{T} \underbrace{\left(x_t^\top \theta_{t-1} - y_t \right)^2 x_t^\top G_t^{-1} x_t}_{\text{first term}}$$

$$\le 4(Y^r)^2 \sum_{t=1}^{T} x_t^\top G_t^{-1} x_t$$

For the forward algorithm:

$$L_T^{\mathrm{f}} - \min_\theta \left(L_T(\theta) + \lambda \|\theta\|_2^2 \right) \le \sum_{t=1}^{T} \underbrace{y_t^2 x_t^\top G_t^{-1} x_t}_{\text{first term}} - \sum_{t=1}^{T-1} \underbrace{\boldsymbol{x}_{t+1}^\top \boldsymbol{G}_t^{-1} \boldsymbol{x}_{t+1} \left(\boldsymbol{x}_{t+1}^\top \boldsymbol{\theta}_t \right)^2}_{\text{second term}}$$

$$\le Y^2 \sum_{t=1}^{T} x_t^\top G_t^{-1} x_t + 0$$

Limitations of the analysis

1/2: Too generic

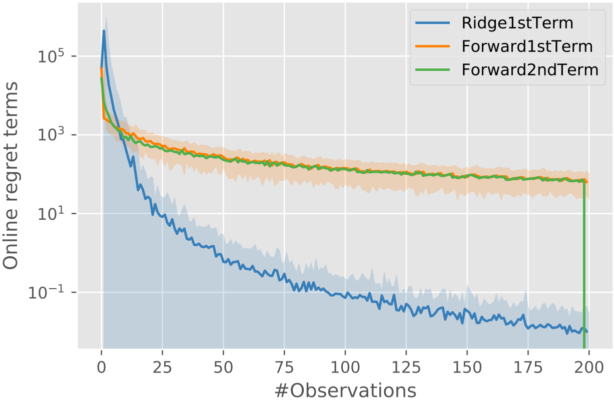


Figure: Online regret. y-axis is logarithmic.

Limitations of the analysis

2/2: bounded observations

Time dependence: $Y^{\mathcal{A}}$ hides a time-dependence

- For the forward algorithm: $Y^f = \max_{1 \le t \le T} |x_t^\top \theta_* + \epsilon_t|$ verifies

$$\forall T \geq 1 : \mathbb{E}[Y^{\mathrm{f}}] \geq \mathbb{E}\left[\max_{1 \le t \le T} \epsilon_t\right] - X\|\theta_*\|_2 \geq \sigma C \sqrt{2\log(T)} - X\|\theta_*\|_2$$

- $(Y^{\mathcal{A}})^2$ appears in previous bounds $\implies$ stochastic regret of order $\log(T)^2$.

Is the sub-Gaussian stochastic setting strictly harder than the bounded adversarial counterpart?

High probability bounds (contributions)

Stochastic setting, sub-Gaussian noise

Recall the regret $R_T^{\mathcal{A}} = L_T^{\mathcal{A}} - \min_{\theta'} L_T(\theta')$.

Definition (new regret) for algo $\mathcal{A}$, define $\bar{R}_T^{\mathcal{A}} = L_T^{\mathcal{A}} - L_T(\theta_*)$.

Theorem (Regret equivalence) w.p $1-\delta$, for all $\|x_t\|_2 \leq X, |G_T(0)| > 0$

$$R_T^{\mathcal{A}} = \bar{R}_T^{\mathcal{A}} + o\left(\log(T)^2\right)$$

High probability bounds (contributions)

Stochastic setting, sub-Gaussian noise

Recall the regret $R_T^{\mathcal{A}} = L_T^{\mathcal{A}} - \min_{\theta'} L_T(\theta')$.

Definition (new regret) for algo $\mathcal{A}$, define $\bar{R}_T^{\mathcal{A}} = L_T^{\mathcal{A}} - L_T(\theta_*)$.

Theorem (Regret equivalence) w.p $1 - \delta$, for all $\|x_t\|_2 \leq X, |G_T(0)| > 0$

$$R_T^{\mathcal{A}} = \bar{R}_T^{\mathcal{A}} + o\left(\log(T)^2\right)$$

Remarks:

- This equivalence is hard to prove because it doesn't involve a regularization, we needed uniform confidence intervals that hold once the design matrix is non-singular
- We will see that the regret is of order $\log(T)^2$, thus the previous result implies that we can use the new regret definition equivalently
- Proof sketch: the difference of the two regrets is basically due to the noise, we control the latter with concentration arguments

High probability bounds (contributions)

Stochastic setting, sub-Gaussian noise

Theorem (Regret bounds) w.p $1-\delta$,

$$\bar{R}_T^{\mathtt{r}} \leq \frac{2d\sigma^2 X^2}{\lambda \log(1+X^2/\lambda)} \log(T) \log\left(T^{d/2}/\delta\right) + o(\log(T)^2),$$

$$\bar{R}_T^{\mathtt{f}} \leq 2d\sigma^2 \log(T) \log\left(T^{d/2}/\delta\right) + o(\log(T)^2),$$

where $X = \max_{1 \leq t \leq T} \|x_t\|_2$.

High probability bounds (contributions)

Stochastic setting, sub-Gaussian noise

Theorem (Regret bounds) w.p $1-\delta$,

$$\bar{R}_T^{\mathrm{r}} \leq \frac{2d\sigma^2 X^2}{\lambda \log(1+X^2/\lambda)} \log(T) \log\left(T^{d/2}/\delta\right) + o(\log(T)^2),$$

$$\bar{R}_T^{\mathrm{f}} \leq 2d\sigma^2 \log(T) \log\left(T^{d/2}/\delta\right) + o(\log(T)^2),$$

where $X = \max_{1\leq t\leq T} \|x_t\|_2$.

Remarks

- The bounds hold with high probability uniformly over T
- For Ridge regression, $1/\lambda$ emerges from bounding $1/\lambda_{min}(G_t(0))$ in the worst case.
- This analysis lifts the "stringent regularization" that requires $\lambda = 1/T$. Therefore, our theorems are not a mere consequence of bounding Y^2 with high probability in previous deterministic theorems.

Tightness of our bounds
1/2

Theorem [Tirinzoni et al., 2020] Let $\delta \in (0,1)$, $n \geq 3$, and $\widehat{\theta}_t$ be a regularized least-square estimator obtained using $t \in [n]$ samples, then

$$\mathbb{P}\left\{\exists t \in [n] : \left\|\widehat{\theta}_t - \theta_*\right\|_{\bar{V}_t} \geq \sqrt{c_{n,\delta}}\right\} \leq \delta$$

where $c_{n,\delta}$ is of order $\mathcal{O}(\log(1/\delta) + d \log\log n)$.

Implication:

Improved upper-bound: $R_T = O\left(d\sigma^2 \log(T) \log\log(T)\right)$.

Tightness of our bounds
2/2

Theorem [Mourtada, 2019], Informal: The worst case online regret scales as:

$$\inf_{\widehat{\beta}_n} \sup_{P \in \mathcal{P}_{\text{Gauss}}(P_X, \sigma^2)} \mathbb{E}\left[\mathcal{E}_P\left(\widehat{\beta}_n\right)\right] \geqslant \frac{\sigma^2 d}{n - d + 1}$$

Implications:

- The regret lower bound is of order: $\mathcal{O}(\sigma^2 d \log(n))$
- The improved bounds are of order $\tilde{\mathcal{O}}(\sigma^2 d \log(n))$ where the $\tilde{\mathcal{O}}$ hides sub-logarithmic factors
- This suggests *the optimality of the forward algorithm*

Experiments

Empirical comparison

Experiment description: We consider a 5 dimensional regression setting, and we vary $\lambda \in \{1/T, 1/\log(T), 1, 10\}$. The noise is drawn from a Gaussian with $\sigma = 0.1$ and features are drawn randomly from the unit ball.

Experiments

Empirical comparison

Experiment description: We consider a 5 dimensional regression setting, and we vary $\lambda \in \{1/T, 1/\log(T), 1, 10\}$. The noise is drawn from a Gaussian with $\sigma = 0.1$ and features are drawn randomly from the unit ball.

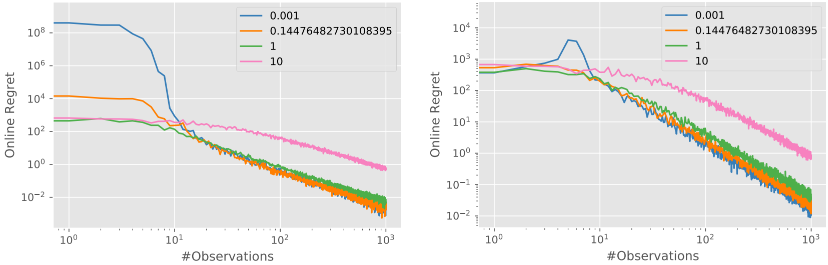


Figure: Left is Ridge regression and right is the Forward algorithm. **Axes are logarithmic**. Performance is averaged over 100 repetitions and shaded areas represent one standard deviation.

Application: linear bandits

Setting

Interaction protocol: At step t

- Nature provides the action space $\mathcal{X}_t \subset \mathbb{R}^d$
- Learner chooses an action (arm) $x_t \in \mathcal{X}_t$
- Nature reveals reward for the selected arm

Application: linear bandits

Setting

Interaction protocol: At step t

- Nature provides the action space $\mathcal{X}_t \subset \mathbb{R}^d$
- Learner chooses an action (arm) $x_t \in \mathcal{X}_t$
- Nature reveals reward for the selected arm

Linear setting: In linear bandits, the reward of action x_t at time t is

$$y_t = \langle x_t, \theta_* \rangle + \epsilon_t$$

where $\theta_* \in \mathbb{R}^d$ is unknown, $\|\theta_*\|_2 \leq S$.

Objective: The learner aims to minimize the (pseudo)regret defined as:

$$R_T = \max_{x \in \mathcal{X}} \sum_{t=1}^{T} \langle x - x_t, \theta_* \rangle$$

Application: linear bandits

Standard algorithm: 1/2

Optimism in the face of uncertainty [Abbasi-Yadkori et al., 2011]: Ridge regression and choose arm that maximizes the upper confidence bound.

Algorithm 1 OFUL algorithm

Input parameters: $\lambda, \delta, S > 0$
for all $t = 1, \ldots, T$
    Define $G_{t-1,x} = \sum_{s=1}^{t-1} x_s x_s^\top$
    Define $\theta_t^r = \text{argmin}_{\theta \in \mathbb{R}^d, \|\theta\|_2 \le S} \sum_{s=1}^{t-1} (y_s - \langle x_s, \theta \rangle)^2 + \lambda \|\theta\|_2^2$
    Define $x_t = \text{argmax}_{x \in \mathcal{X}} \langle x, \theta_t^r \rangle + \|x\|_{G_{t-1}^{-1}} (1 + S\sqrt{\lambda}) + \sigma \sqrt{2 \log \left(\frac{(1 + tX^2/\lambda d)^{d/2}}{\delta} \right)}$
    play x_t and observe y_t.
end for

Application: linear bandits

Standard algorithm: 2/2

Standard assumption: for all $x_t \in \mathcal{X} \quad \langle x_t, \theta_* \rangle \in [-1, 1]$.

Application: linear bandits

Standard algorithm: 2/2

Standard assumption: for all $x_t \in \mathcal{X} \quad \langle x_t, \theta_* \rangle \in [-1, 1]$.

Theorem [Abbasi-Yadkori et al., 2011] Under the above assumption, w.p $1-\delta$, for all $T > 0$:

$$R_T^{\mathrm{r}} \leq 4\sqrt{Td\log(\lambda + TX^2/d)}\left(\sigma\sqrt{2\log(1/\delta) + d\log(1 + TX^2/(\lambda d))} + S\sqrt{\lambda}\right),$$

where $X = \max_{1 \leq t \leq T} \|x_t\|_2$.

Application: linear bandits

Standard algorithm: 2/2

Standard assumption: for all $x_t \in \mathcal{X} \quad \langle x_t, \theta_* \rangle \in [-1, 1]\,.$

Theorem [Abbasi-Yadkori et al., 2011] Under the above assumption, w.p $1 - \delta$, for all $T > 0$:

$$R_T^{\mathrm{r}} \leq 4\sqrt{Td \log(\lambda + TX^2/d)}\left(\sigma\sqrt{2\log(1/\delta) + d\log(1 + TX^2/(\lambda d))} + S\sqrt{\lambda}\right),$$

where $X = \max_{1 \leq t \leq T} \|x_t\|_2$.

Limitations:

- The algorithm requires prior knowledge of the rewards' bound for the confidence bound construction
- If we study the regret instead of pseudo-regret, the standard analysis no longer holds

Application: linear bandits

Contribution: improved OFUL algorithm

OFUL^f: Forward regression mixed with UCB.

Algorithm 2 OFUL$^{\mathrm{f}}$ algorithm

Input parameters: $\lambda, \delta, S > 0$

for all $t = 1, \ldots, T$

Define $X_t(x) = \max\{\|x\|_2, \max_{1 \le s \le t-1} \|x_s\|_2\}$, $G_{t-1,x} = G_{t-1} + xx^\top$

Define $\theta_t^{\mathrm{f}}(x) = \operatorname{argmin}_{\theta \in \mathbb{R}^d} \sum_{s=1}^{t-1} (y_s - \langle x_s, \theta \rangle)^2 + \lambda \|\theta\|_2^2 + \langle x, \theta \rangle^2$

Define $x_t = \operatorname{argmax}_{x \in \mathcal{X}} \langle x, \theta_t^{\mathrm{f}}(x) \rangle + \|x\|_{G_{t-1,x}^{-1}} (\sqrt{\lambda} + \|x\|_2) S$

$$+ \sigma \sqrt{2 \log \left(\frac{(1 + t X_t^2(x)/\lambda d)^{d/2}}{\delta} \right)}$$

play x_t and observe y_t.

end for

Application: linear bandits

OFUL$^{\mathrm{f}}$: analysis

Theorem: w.p $1-\delta$, for all $T \geq 1$:

$$R_T^{\mathrm{f}} \leq 4\sqrt{Td\log(\lambda + TX^2/d)}\left(\sigma\sqrt{2\log(1/\delta) + d\log(1 + TX^2/(\lambda d))} + (\lambda^{1/2} + X)S\right).$$

Application: linear bandits

OFUL$^{\mathrm{f}}$: analysis

Theorem: w.p $1-\delta$, for all $T \geq 1$:

$$R_T^{\mathrm{f}} \leq 4\sqrt{Td\log(\lambda + TX^2/d)}\left(\sigma\sqrt{2\log(1/\delta) + d\log(1+TX^2/(\lambda d))} + (\lambda^{1/2} + X)S\right).$$

Without the common assumption, w.p $1-\delta$, for all $T \geq 1$,

$$R_T^{\mathrm{r}} \leq 4\sqrt{\frac{\boldsymbol{X}^2\, Td\log(\lambda + TX^2/d)}{\boldsymbol{\lambda}\log(1+\boldsymbol{X}^2/\boldsymbol{\lambda})}}\left(S\sqrt{\lambda} + \sigma\sqrt{2\log(1/\delta) + d\log(1+TX^2/(\lambda d))}\right)$$

Application: linear bandits

Experiment: description

Setup: Consider a 100-dimensional linear bandit with 10 arms:

- θ_* is drawn from the unit ball
- Actions are such that $\|x_t\|_2 \leq 200$
- Noise $\epsilon_t \overset{\mathcal{L}}{=} \mathcal{N}(0, 10^{-1})$, $T = 10^5$, $\lambda = 10^{-5}$, $\delta = 10^{-3}$.

Regularization choice We choose $\lambda = 1/T$, there are two reasons for this:

1. The adversarial bounds suggest that $\lambda = 1/T$ is best, and we want to demonstrate the benefits of our stochastic analysis
2. To showcase the increased robustness of OFUL^f compared to OFUL. (More often than not, OFUL performs as good as OFUL^f)

Application: linear bandits

Experiment: result

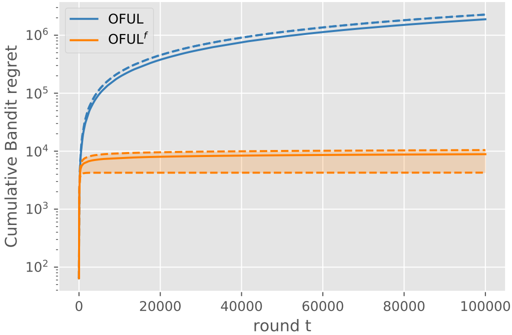


Figure: Cumulative regret. y-axis is logarithmic.

Observations:

- As predicted by the bounds, $\lambda = 1/T$ incurs linear regret for OFUL
- OFUL$^{\mathtt{f}}$ is robust even for $\lambda = 1/T$
- OFUL$^{\mathtt{f}}$ is robust and drops the bounded rewards assumption

Unregularized forward

The unregularized forward algorithm [Gaillard et al., 2019]

Input: initial $\theta_0 \in \mathbb{R}^d$, and regularization $\lambda \in \mathbb{R}_+$
for $t = 1, 2, \ldots, T$
  Observe input x_t
  $\theta^{uf}_{t-1} \in \text{argmin}_\theta \, L_{t-1}(\theta)$ $+(x_t^T \theta)^2$
  predict $\hat{y}_t = x_t^T \theta^{uf}_{t-1} \in \mathbb{R}$
  observe y_t
  incur $\ell_t \in \mathbb{R}$
end for

One of the solutions of the optimization has a closed form:

$$\theta^f_t = G^\dagger_{t+1} b_t = (\underbrace{G_t}_{\text{Same as Ridge}} + x_{t+1} x_{t+1}^\top)^\dagger \underbrace{b_t}_{\text{Same as Ridge}}$$

Unregularized forward

Existing results: adversarial setting

The old bounds are not usable: they blow up for $\lambda = 0$

Theorem [Gaillard et al., 2019] For all $T \geq 1$, for all sequences $\mathbf{x}_1, \dots, \mathbf{x}_T \in \mathbb{R}^d$ such that $\|x_t\|_2 \leq X$ and all $y_1, \dots, y_T \in [-Y, Y]$, the unregularized forward achieves the uniform regret bound

$$\sup_{\mathbf{u} \in \mathbb{R}^d} \mathcal{R}_T(\mathbf{u}) \leqslant dY^2 \ln T + dY^2 + Y^2 \sum_{t \in [1,T] \cap \mathcal{T}} \ln \left(\frac{X^2}{\lambda_{r_t}(\mathbf{G}_t)} \right).$$

where the set $\mathcal{T}$ contains r_T rounds for which $\operatorname{rank}(\mathbf{G}_{s-1}) \neq \operatorname{rank}(\mathbf{G}_s)$.

This algorithm is optimal: it matches the known lower bound, but the second term could be (and stay) arbitrarily large

Unregularized forward

Stochastic setting: contribution

Theorem The unregularized forward regression achieves, for any $\delta > 0$, with probability at least $1-\delta$ for all $T > 0$:

$$\bar{R}_T^{u-f} \le 2(1+\kappa)(1+\alpha)\sigma^2 \log\left(\frac{\kappa_d(1+TX^2/\gamma d)}{\delta/4}\right)\log\left(\frac{|G_T^\dagger|}{|G_{T_1}^\dagger|}\right)$$

$$+ 2\sigma^2 \log\left(\frac{4T_1}{\delta}\right)\left(d + \sum_{1\le t\le T_1, t\in\mathcal{T}} \log\left(\frac{X^2}{\lambda_{r_t}(\sum_{s=1}^t x_t x_t^\top)}\right)\right),$$

where $\kappa, \alpha \in \mathbb{R}_+^*$ are (chosen) peeling parameters, $\gamma = \min_{1\le t\le T}\|x_t\|_2$, . $T_1 = \min\{t \ge 1, |G_t| > 0\}$ is the first time the design matrix is non-singular.

Conclusion and perspectives

This paper:

◦ Revisits the analysis of online linear regression algorithms in the stochastic setup, with possibly unbounded observations

◦ Provides novel understanding of online regression algorithms:

1. The first analysis of ridge regression without bounded predictions or prior knowledge of the observations' bound
2. Our novel bounds seem to correctly capture the dependence with regularization

◦ Replaces ridge by forward in linear approximations:

1. Theory: forward enjoys standard regret; drops boundedness Assumption
2. Practice: forward makes the algorithm robust to regularization

◦ Claims that the forward improvement is expandable: we provide the analysis for non-stationary linear bandits in the appendix of our paper

Thank you for your interest in the paper

Questions?

Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. (2011). Improved algorithms for linear stochastic bandits.
In Advances in Neural Information Processing Systems, pages 2312–2320.

Azoury, K. S. and Warmuth, M. K. (2001).
Relative loss bounds for on-line density estimation with the exponential family of distributions.
Machine Learning, 43(3):211–246.

Gaillard, P., Gerchinovitz, S., Huard, M., and Stoltz, G. (2019).
Uniform regret bounds over $\mathbb{R}^d$ for the sequential linear regression problem with the square loss.
In Algorithmic Learning Theory, pages 404–432.

Mourtada, J. (2019).
Exact minimax risk for linear least squares, and the lower tail of sample covariance matrices.
arXiv preprint arXiv:1912.10754.

Tirinzoni, A., Pirotta, M., Restelli, M., and Lazaric, A. (2020).
An asymptotically optimal primal-dual incremental algorithm for contextual linear bandits.
arXiv preprint arXiv:2010.12247.

Vovk, V. (2001).
Competitive on-line statistics.
International Statistical Review, 69(2):213–248.