Stochastic Online Linear Regression: the Forward Algorithm to Replace Ridge

Reda Ouhamma*, Odalric Maillard*, Vianney Perchet†

* Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France. † Crest, Ensae & Criteo AI Lab

NEURAL INFORMATION PROCESSING SYSTEMS

Setting

Nature chooses features, parameter $(x_t)_{t\geq 1}, \theta_* \in \mathbb{R}^d$, $\forall t\ y_t = x_t^\top \theta_* + \epsilon_t$.

At step t, the learner:

- observes $x_t \in \mathbb{R}^d$
- chooses $\theta_{t-1} \in \mathbb{R}^d$ and predicts $\hat{y}_t = x_t^T \theta_{t-1}$
- observes label y_t and suffers loss $\ell_t = (y_t - \hat{y}_t)^2$

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

→ In the following, we are interested in two popular algorithms:

Algorithm 1 Online ridge regression

initialize $\theta_0 \in \mathbb{R}^d, \lambda \in \mathbb{R}_+$
for $t = 1, 2, \dots, T$ **do**
  Observe input x_t
  predict $\hat{y}_t = x_t^T \theta_{t-1}^r \in \mathbb{R}$
  observe y_t
  incur $\ell_t \in \mathbb{R}$
  $\theta_t^r \in \arg\min_\theta L_t(\theta) + \lambda\|\theta\|_2^2$
end for

Algorithm 2 Forward regression

initialize $\theta_0 \in \mathbb{R}^d, \lambda \in \mathbb{R}_+$
for $t = 1, 2, \dots, T$ **do**
  Observe input x_t
  $\theta_{t-1}^f \in \arg\min_\theta L_{t-1}(\theta) + \lambda\|\theta\|_2^2 + (x_t^T\theta)^2$
  predict $\hat{y}_t = x_t^T \theta_{t-1}^f \in \mathbb{R}$
  observe y_t
  incur $\ell_t \in \mathbb{R}$
end for

Existing analysis and limitations

Consider $\underline{R}_T = L_T - \min_\theta \left(L_T(\theta) + \lambda\|\theta\|_2^2\right)$. If the observations y_t are adversarial, then it can be shown that:

$$\underline{R}_T^r \leq 4\left(Y^r\right)^2 d\log\left(1 + \frac{TX^2}{\lambda d}\right), \quad \underline{R}_T^f \leq \left(Y^f\right)^2 d\log\left(1 + \frac{TX^2}{\lambda d}\right),$$

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

Adversarial setting: previous bounds suffer from a rigid regularization: constraint: we can write

$$R_T^{\mathcal{A}} \leq c^{\mathcal{A}}\left(Y^{\mathcal{A}}\right)^2 d\log\left(1 + \frac{TX^2}{\lambda d}\right) + \frac{\lambda\left(Y^{\mathcal{A}}\right)^2 T}{\lambda_{r_T}\left(G_T(0)\right)}$$

where λ_{r_T} is the smallest positive eigenvalue of the design matrix. This forces the choice $\lambda \sim 1/T$ to obtain a logarithmic regret.

Stochastic setting: existing bounds suffer from two limitations:

1. *Too loose*, the adversarial analysis proceeds as follows:
   - $\underline{R}_T^r \leq \sum_{t=1}^T \left(x_t^\top \theta_{t-1} - y_t\right)^2 x_t^\top G_t^{-1} x_t \leq 4(Y^r)^2 \sum_{t=1}^T x_t^\top G_t^{-1} x_t$. while in a stochastic setting, $\left(x_t^\top \theta_{t-1} - y_t\right)^2 \sim \epsilon_t^2 \ll (Y^r)^2$.
   - $\underline{R}_T^f \leq \sum_{t=1}^T y_t^2 x_t^\top G_t^{-1} x_t - \sum_{t=1}^{T-1} \boldsymbol{x}_{t+1}^\top \boldsymbol{G}_t^{-1} \boldsymbol{x}_{t+1}\left(\boldsymbol{x}_{t+1}^\top \boldsymbol{\theta}_t\right)^2$. where the second term is neglected to conclude.
2. *Time dependence:* in a stochastic setting, $Y^{\mathcal{A}}$ hides a time-dependence, for all $T \geq 1$:

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

High probability bounds

Definition: we use a new regret definition that is more convenient for the setting. For an algo $\mathcal{A}$, define $\quad \bar{R}_T^{\mathcal{A}} = L_T^{\mathcal{A}} - L_T(\theta_*)$.

Theorem (Regret equivalence) If $|G_T(0)| > 0$, with probability $1-\delta$, for all $\|x_t\|_2 \leq X$, $\quad R_T^{\mathcal{A}} = \bar{R}_T^{\mathcal{A}} + o\left(\log(T)^2\right)$

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

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

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

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

Remarks:

- The range of predictions and observations no longer appears
- A factor $1/\lambda$ appears for ridge regression, it is the price for not assuming bounded predictions
- Both results lift the stringent regularization condition

Tightness of the bounds: a tighter concentration for regularized least squares [1] was derived concurrently with the writing of this paper:

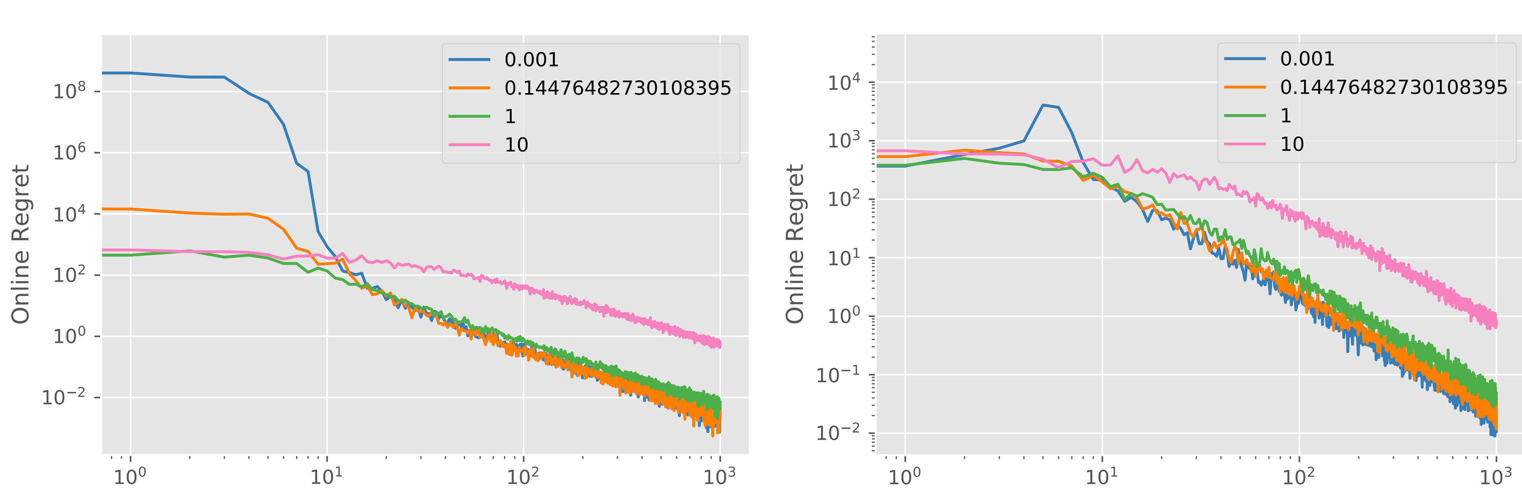
- They improve the width of the confidence interval from $d\log(T/\delta)$ to $\log(1/\delta) + d\log\log(T)$.
- Injecting in our proof gives

$$R_T = O\left(d\sigma^2\log(T)\log(1/\delta) + (d\sigma)^2\log(T)\log\log(T)\right)$$

- The latter matches Theorem 1 in [2] up to $\log(1/\delta) + d\log\log(T)$ suggesting that the forward algorithm is nearly optimal.

Experiments

We provide experimental evidence supporting the fact that our novel high probability analysis better reflects the influence of regularization than results its adversarial counterpart.



Dependence on λ, axes are logarithmic. Left is ridge regression and right is forward regression, performances are averaged over 100 realizations.

Comment: once the design matrix $G_t(0)$ becomes non-singular, the $1/\lambda$ virtually disappears from regret bound and is replaced by the smallest eigenvalue of $G_t(0)$, making the regret significantly more stable.

Application: linear bandits

Setting: linear bandits, reward of action x_t at time t is $y_t = \langle x_t, \theta_*\rangle + \epsilon_t$, $\|\theta_*\|_2 \leq S$. The (pseudo) regret is: $R_T = \max_{x\in\mathcal{X}}\sum_{t=1}^T\langle x - x_t, \theta_*\rangle$.

We propose a variant of the OFUL algorithm of [3] using forward.

Algorithm 3 OFUL$^{\mathrm{f}}$ algorithm

Input parameters: $\lambda, \delta, S > 0$
for all $t = 1, \dots, T$ **do**
  Define $X_t(x) = \max\{\|x\|_2, \max_{1\leq s\leq t-1}\|x_s\|_2\}$, $G_{t-1,x} = G_{t-1} + xx^\top$
  Define $\theta_t^{\mathrm{f}}(x) = \arg\min_{\theta\in\mathbb{R}^d}\sum_{s=1}^{t-1}(y_s - \langle x_s, \theta\rangle)^2 + \lambda\|\theta\|_2^2 + \langle x, \theta\rangle^2$
  Define $x_t = \arg\max_{x\in\mathcal{X}}\langle x, \theta_t^{\mathrm{f}}(x)\rangle + \|x\|_{G_{t-1,x}^{-1}}(\sqrt{\lambda} + \|x\|_2)S + \sigma\sqrt{2\log\left(\frac{(1+tX_t^2(x)/\lambda d)^{d/2}}{\delta}\right)}$
  Play x_t and observe y_t.
end for

OFUL requires a strong assumption for the analysis: $|\langle x_t, \theta_*\rangle| \leq 1$. We present a new analysis for OFUL and OFUL^f without it:

Theorem (Regret bounds) with probability at least $1-\delta$, for all $T \geq 1$:

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

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

Remarks:

- $\lambda = 1/T$ incurs linear regret for OFUL OFUL$^{\mathrm{f}}$ is robust
- OFUL$^{\mathrm{f}}$ drops the prior knowledge of bounds assumption
- OFUL$^{\mathrm{f}}$ is significantly more robust in practice

Conclusion

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. First ridge regret bound regression without bounded predictions or knowledge 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

References

[1] Andrea Tirinzoni, Matteo Pirotta, Marcello Restelli, and Alessandro Lazaric. An asymptotically optimal primal-dual incremental algorithm for contextual linear bandits. *arXiv preprint arXiv:2010.12247*, 2020.

[2] Jaouad Mourtada. Exact minimax risk for linear least squares, and the lower tail of sample covariance matrices. *arXiv preprint arXiv:1912.10754*, 2019.

[3] Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. *Advances in neural information processing systems*, 24:2312–2320, 2011.