CSC 665: Online to batch conversion

Chicheng Zhang

December 10, 2019

1 A slight reformulation of online learning

High level idea: suppose we have a online learning algorithm that has a regret guarantee against *any* sequences of examples, then when it is run on a sequence of examples iid from D, the intermediate predictors it outputs will have good statistical guarantees. Specifically, we will show that the *average excess loss* of these predictors will be small if the regret is small.

First, let us describe a simple reformulation of online learning. Instead of asking the learner to make a prediction after seeing the context information of an example, we directly ask the learner to output a predictor. These two formulations are equivalent (the latter is called skolemization), in that predictor h_t at round t can be thought of as a strategy to respond to every possible context x_t at that round.

For example, in online classification where $\mathcal{A} = \{\pm 1\}$, h_t is a binary classifier, and $\ell_t(h) = \mathbf{1}(h(x_t) \neq y_t)$ is the mistake indicator of classifier h on example (x_t, y_t).

Algorithm 1 Online learning: reformulation

Require: Context space $\mathcal{X}$, action space $\mathcal{A}$, hypothesis class $\mathcal{H}$.
  for timesteps $t = 1, 2, \ldots, T$: **do**
    Show predictor $h_t : \mathcal{X} \to \mathcal{A}$.
    Receive loss ℓ_t, where for every predictor h, $\ell_t(h) \in [0, M]$ is the instaneous loss of h at round t.
  end for
  Goal: minimize cumulative regret $\mathrm{Reg}(T, \mathcal{H}) = \sum_{t=1}^T \ell_t(h_t) - \min_{h \in \mathcal{H}} \sum_{t=1}^T \ell_t(h_t)$.

2 Online to batch conversion

Suppose that we have an online learning algorithm $\mathcal{A}$ that achieves a small cumulative regret, that is, $\mathrm{Reg}(T, \mathcal{H}) \leq \mathrm{R}(T)$; for example, if $\mathcal{H}$ is finite, one can run (variants of) Hedge on $\mathcal{H}$ to get a regret guarantee $\mathrm{R}(T) = O(M\sqrt{\ln |\mathcal{H}| T})$.

Let us consider the statistical learning setting, where we have a distribution D over loss functions. To evaluate the performance of a predictor h on distribution D, we use the familiar expected loss, that is, $L(h, D) = \mathbb{E}_{\ell \sim D}\, \ell(h)$.

Consider running algorithm $\mathcal{A}$, with losses ℓ_t iid from D. We would like to show that, the classifiers $\{h_1, \ldots, h_T\}$ generated by $\mathcal{A}$ has a low average excess loss with respect to D. Specifically, we have the following theorem.

Theorem 1 ([1]). *Suppose the setting is as described above. Then,*

$$\frac{1}{T} \mathbb{E}\left[\sum_{t=1}^T L(h_t, D)\right] \leq \min_{h' \in \mathcal{H}} \mathbb{E}[L(h', D)] + \frac{\mathrm{R}(T)}{T}. \quad (1)$$

where the expectation is wrt the randomness of both the algorithm $\mathcal{A}$ and the losses drawn from D. Furthermore, with probability $1-\delta$,

$$\frac{1}{T}\,\mathbb{E}\left[\sum_{t=1}^{T} L(h_t, D)\right] \le \min_{h'\in\mathcal{H}} \mathbb{E}[L(h', D)] + \frac{\mathrm{R}(T)}{T} + M\sqrt{\frac{8\ln\frac{4}{\delta}}{T}}. \tag{2}$$

Remark. In many online learning settings, $\mathrm{R}(T)$ is $\Omega(\sqrt{T})$, and therefore the extra $\sqrt{\frac{1}{T}}$ factor paid in concentration is dominated by $\frac{\mathrm{R}(T)}{T}$. If $\mathrm{R}(T)$'s are of smaller order, then this theorem no longer provides the tightest conversion result; in these settings, more advanced concentration inequalities such as Bernstein's and Freedman's Inequalities provide better results; but this is beyond the scope of this course.

Proof. Let us first write down the online regret guarantee on loss sequence $\{\ell_t\}_{t=1}^{T}$:

$$\sum_{t=1}^{T} \ell_t(h_t) \le \min_{h'\in\mathcal{H}} \sum_{t=1}^{T} \ell_t(h') + \mathrm{R}(T). \tag{3}$$

Proof of Equation (1)**.** Both sides of the inequality involves random variables. As the above inequality holds for any random seed, the expectation of the left hand side is also at most the expectation of the right hand side. By linearity of expectation, $\mathbb{E}\sum_{t=1}^{T}\ell_t(h_t) = \sum_{t=1}^{T}\mathbb{E}\,\ell_t(h_t)$. Now, define H_t as the collection of random variables up to $t-1$, along with h_t, that is $H_t = (h_1, \ell_1, \ldots, h_{t-1}, \ell_{t-1}, h_t)$.

Note that conditioned on H_t, the only randomness in $\ell_t(h_t)$ comes from the fresh loss ℓ_t. Consequently,

$$\mathbb{E}\left[\ell_t(h_t)|H_t\right] = \mathbb{E}_{\ell\sim D}\,\ell(h_t) = L(h_t, D).$$

By the law of iterated expectation, $\mathbb{E}[\ell_t(h_t)] = \mathbb{E}\left[\mathbb{E}\left[\ell_t(h_t)|H_t\right]\right] = \mathbb{E}[L(h_t, D)]$.

Therefore, the expectation of the left hand side is $\sum_{t=1}^{T}\mathbb{E}[\sum L(h_t, D)]$.

What's the expectation of the first term on the right hand side? In fact we have seen this in statistical learning. Note that $f(x) = \min(x_1, \ldots, x_n)$ is a concave function. By Jensen's Inequality,

$$\mathbb{E}\left[\min_{h'\in\mathcal{H}}\sum_{t=1}^{T}\ell_t(h')\right] \le \min_{h'\in\mathcal{H}}\mathbb{E}[\sum_{t=1}^{T}\ell_t(h')] = T\min_{h'\in\mathcal{H}} L(h', D).$$

Equation (1) is shown by combining the above observations, taking expectations, and dividing both sides by T.

Proof of Equation (2)**.** Define $X_t = \ell_t(h_t) - L(h_t, D)$. Observe that by the above discussion, $\{X_t\}_{t=1}^{T}$ is a martingale difference sequence with respect to history $\{H_t\}_{t=1}^{T}$. Moreover, $|X_t| \le M$, by our assumption that $\ell_t(h) \in [0, M]$ for all t and h. Therefore, with probability $1-\delta/2$,

$$\left|\sum_{t=1}^{T} X_t\right| = \sum_{t=1}^{T}\ell_t(h_t) - \sum_{t=1}^{T} L(h_t, D) \le M\sqrt{2T\ln\frac{4}{\delta}}. \tag{4}$$

In addition, we can upper bound $\min_{h'\in\mathcal{H}}\sum_{t=1}^{T}\ell_t(h')$ as follows. Denote by $h^\star = \mathrm{argmin}_{h\in\mathcal{H}} L(h, D)$. We have that with probability $1-\delta/2$,

$$\sum_{t=1}^{T}\ell_t(h^\star) \le T L(h^\star, D) + M\sqrt{2T\ln\frac{4}{\delta}}. \tag{5}$$

Consequently,

$$\begin{aligned}\min_{h'\in\mathcal{H}}\sum_{t=1}^T \ell_t(h') &\leq \sum_{t=1}^T \ell_t(h^\star) \\ &\leq TL(h^\star,D) + M\sqrt{2T\ln\frac{4}{\delta}} \\ &= T\min_{h'\in\mathcal{H}} L(h'^\star,D) + M\sqrt{2T\ln\frac{4}{\delta}} \qquad (6)\end{aligned}$$

holds with probability $1-\delta/2$. Here the first inequality is from the suboptimality of $h^\star$; the second inequality is from Equation (5); the third inequality uses the definition of $h^\star$.

Now, combining Equations (3), (4), (6) with union bound, and divide both sides by T, we get Equation (2) happens with probability $1-\delta$. □

Remark. There are two ways of utilizing the guarantee in Equation (2) by returning a predictor $\tilde{h}$ that has excess loss guarantees:

1. Let $\tilde{h}$ be an element chosen uniformly at random from $\{h_1,\ldots,h_T\}$. Note that we have

$$\mathbb{E}_{\tilde{h}\sim \mathrm{U}(\{h_1,\ldots,h_T\})} L(\tilde{h},D) = \frac{1}{T}\sum_{t=1}^T L(h_t,D).$$

   Therefore, the expected loss of $\tilde{h}$ has the same upper bound as $\frac{1}{T}\sum_{t=1}^T L(h_t,D)$ in (2).

2. Suppose that $\ell(h)$ is convex with respect to h with probability 1 for $\ell\sim D$; specifically, $\ell(\alpha h_1 + (1-\alpha)h_2) \leq \alpha\ell(h_1) + (1-\alpha)\ell(h_2)$ for all h_1,h_2 and $\alpha\in[0,1]$; for example, $\ell(h) = (h(x)-y)^2$ or $\ell_t(h) = |h(x)-y|$. Denote by $\bar{h} = \frac{1}{T}\sum_{t=1}^T h_t$. By Jensen's Inequality, we have:

$$L(\bar{h},D) \leq \frac{1}{T}\sum_{t=1}^T L(h_t,D).$$

   Therefore, the expected loss of $\bar{h}$ has the same upper bound as $\frac{1}{T}\sum_{t=1}^T L(h_t,D)$ in (2).

References

[1] Nicolo Cesa-Bianchi, Alex Conconi, and Claudio Gentile. On the generalization ability of on-line learning algorithms. *IEEE Transactions on Information Theory*, 50(9):2050–2057, 2004.

A Azuma-Hoeffding's Inequality

The fact below has many applications for sequential decision applications. We consider a martingale difference sequence, a sequence of random variables that are only "weakly dependent", and show that they exhibit concentration behavior similar to independent random variables.

Definition 1. *$X_1,\ldots,X_T$ is a martingale difference sequence with respect to history $H_1,\ldots,H_T$, if:*

1. *H_t's are nested: $H_1\subset H_2\subset\ldots\subset H_T$.*
2. *X_t can be written as a function of H_t.*
3. $\mathbb{E}\left[X_t|H_{t-1}\right] = 0$.

Remark. There is a more generation definition of martingale difference sequence in probability theory, defined using filtrations and σ-algebras. We encourage the interested reader to consult standard probability theory textbooks.

It can be checked that a set of iid zero mean random variables is a martingale difference sequence. Moreover, it can be checked that if $X_1, \ldots, X_T$ is a martingale difference sequence, then $\mathbb{E}[X_i X_j] = \mathbb{E}\left[\mathbb{E}\left[X_i | H_j\right] X_j\right] = 0$, i.e. they are pairwise uncorrelated. A canonical example of a martingale difference sequence is the wealth increment at a fair betting game: suppose Z_t's are iid Rademacher random variables, encoding winning $(+1)$ or losing (-1) of a gambler at round t. Let $M_t = f(Z_1, \ldots, Z_{t-1}) \cdot Z_t$, where f encodes the bet the gambler placed at round t. It can be checked that $\{M_t\}_{t=1}^T$ is a martingale difference strategy wrt history $\left\{H_t = (Z_1, \ldots, Z_t)\right\}_{t=1}^T$.

Fact 1 (Azuma-Hoeffding's Inequality). *If $X_1, \ldots, X_T$ is a martingale difference sequence, and for every t, with probability 1, $|X_t| \le M$. Then with probability $1 - \delta$,*

$$\left|\sum_{t=1}^T X_t\right| \le M\sqrt{2T \ln \frac{2}{\delta}}.$$

Note that the above inequality generalizes the classical Hoeffding's Inequality. Intuitively, if a gambler bets adaptively and conservatively (each timestep her wager is at most M), then the typical value of her cumulative wealth change will be around $\pm O(M\sqrt{T})$, as opposed to the maximum possible value $\pm O(MT)$.