

Advanced Machine Learning

Online Learning Basics

MEHRYAR MOHRI

MOHRI@

COURANT INSTITUTE & GOOGLE RESEARCH

Outline

- Prediction with expert advice
- Weighted Majority algorithm (WM)
- Randomized Majority algorithm (RWM)
- Online-to-batch conversion

Motivation

■ PAC learning:

- distribution fixed over time (training and test).
- IID assumption.

■ On-line learning:

- no distributional assumption.
- worst-case analysis (adversarial).
- mixed training and test.
- Performance measure: mistake model, regret.

General Online Setting

- For $t = 1$ to T do
 - receive instance $x_t \in X$.
 - predict $\hat{y}_t \in Y$.
 - receive label $y_t \in Y$.
 - incur loss $L(\hat{y}_t, y_t)$.
- Classification: $Y = \{0, 1\}$, $L(y, y') = |y' - y|$.
- Regression: $Y \subseteq \mathbb{R}$, $L(y, y') = (y' - y)^2$.
- **Objective:** minimize total loss $\sum_{t=1}^T L(\hat{y}_t, y_t)$.

Prediction with Expert Advice

- For $t = 1$ to T do
 - receive instance $x_t \in X$ and advice $y_{t,i} \in Y, i \in [1, N]$.
 - predict $\hat{y}_t \in Y$.
 - receive label $y_t \in Y$.
 - incur loss $L(\hat{y}_t, y_t)$.
- **Objective:** minimize regret, i.e., difference of total loss incurred and that of the best expert,

$$\text{Regret}(T) = \sum_{t=1}^T L(\hat{y}_t, y_t) - \min_{i=1}^N \sum_{t=1}^T L(\hat{y}_{t,i}, y_t).$$

Halving Algorithm

[see (Mitchell, 1997)]

HALVING(H)

```
1    $H_1 \leftarrow H$ 
2   for  $t \leftarrow 1$  to  $T$  do
3       RECEIVE( $x_t$ )
4        $\hat{y}_t \leftarrow \text{MAJORITYVOTE}(H_t, x_t)$ 
5       RECEIVE( $y_t$ )
6       if  $\hat{y}_t \neq y_t$  then
7            $H_{t+1} \leftarrow \{c \in H_t : c(x_t) = y_t\}$ 
8   return  $H_{T+1}$ 
```

Halving Algorithm - Bound

(Littlestone, 1988)

- **Theorem:** Let H be a finite hypothesis set, then the number of mistakes made by the Halving algorithm is bounded as follows:

$$M_{\text{Halving}(H)} \leq \log_2 |H|.$$

- **Proof:** At each mistake, the hypothesis set is reduced at least by half.

Weighted Majority Algorithm

(Littlestone and Warmuth, 1988)

WEIGHTED-MAJORITY(N experts) $\triangleright y_t, y_{t,i} \in \{0, 1\}$.
 $\beta \in [0, 1]$.

```
1  for  $i \leftarrow 1$  to  $N$  do
2       $w_{1,i} \leftarrow 1$ 
3  for  $t \leftarrow 1$  to  $T$  do
4      RECEIVE( $x_t$ )
5       $\hat{y}_t \leftarrow 1_{\sum_{y_{t,i}=1}^N w_t \geq \sum_{y_{t,i}=0}^N w_t}$   $\triangleright$  weighted majority vote
6      RECEIVE( $y_t$ )
7      if  $\hat{y}_t \neq y_t$  then
8          for  $i \leftarrow 1$  to  $N$  do
9              if ( $y_{t,i} \neq y_t$ ) then
10                  $w_{t+1,i} \leftarrow \beta w_{t,i}$ 
11             else  $w_{t+1,i} \leftarrow w_{t,i}$ 
12  return  $\mathbf{w}_{T+1}$ 
```

Weighted Majority - Bound

- **Theorem:** Let m_t be the number of mistakes made by the WM algorithm till time t and m_t^* that of the best expert. Then, for all t ,

$$m_t \leq \frac{\log N + m_t^* \log \frac{1}{\beta}}{\log \frac{2}{1+\beta}}.$$

- Thus, $m_t \leq O(\log N) + \text{constant} \times \text{best expert.}$
- Realizable case: $m_t \leq O(\log N).$
- Halving algorithm: $\beta = 0.$

Weighted Majority - Proof

- Potential: $\Phi_t = \sum_{i=1}^N w_{t,i}$.
- Upper bound: after each error,

$$\Phi_{t+1} \leq [1/2 + 1/2 \beta] \Phi_t = \left[\frac{1 + \beta}{2} \right] \Phi_t.$$

Thus, $\Phi_t \leq \left[\frac{1 + \beta}{2} \right]^{m_t} N.$

- Lower bound: for any expert i , $\Phi_t \geq w_{t,i} = \beta^{m_{t,i}}$.
- Comparison:
 - $\beta^{m_t^*} \leq \left[\frac{1 + \beta}{2} \right]^{m_t} N$
 - $\Rightarrow m_t^* \log \beta \leq \log N + m_t \log \left[\frac{1 + \beta}{2} \right]$
 - $\Rightarrow m_t \log \left[\frac{2}{1 + \beta} \right] \leq \log N + m_t^* \log \frac{1}{\beta}$.

Weighted Majority - Notes

- **Advantage:** remarkable bound requiring no assumption.
- **Disadvantage:** no deterministic algorithm can achieve a regret $R_T = o(T)$ with the binary loss.
 - better guarantee with randomized WM.
 - better guarantee for WM with convex losses.

Exponential Weighted Average

Algorithm:

- weight update: $w_{t+1,i} \leftarrow w_{t,i} e^{-\eta L(\hat{y}_{t,i}, y_t)} = e^{-\eta L_{t,i}}$.
- prediction: $\hat{y}_t = \frac{\sum_{i=1}^N w_{t,i} y_{t,i}}{\sum_{i=1}^N w_{t,i}}$.

total loss incurred by
expert i up to time t

Theorem: assume that L is convex in its first argument and takes values in $[0, 1]$. Then, for any $\eta > 0$ and any sequence $y_1, \dots, y_T \in Y$, the regret at time T satisfies

$$\text{Regret}(T) \leq \frac{\log N}{\eta} + \frac{\eta T}{8}.$$

For $\eta = \sqrt{8 \log N / T}$,

$$\boxed{\text{Regret}(T) \leq \sqrt{(T/2) \log N}}.$$

EW - Proof

■ Potential: $\Phi_t = \log \sum_{i=1}^N w_{t,i}$.

■ Upper bound:

$$\begin{aligned}\Phi_t - \Phi_{t-1} &= \log \frac{\sum_{i=1}^N w_{t-1,i} e^{-\eta L(\hat{y}_{t,i}, y_t)}}{\sum_{i=1}^N w_{t-1,i}} \\ &= \log \left(\underset{w_{t-1}}{\text{E}} [e^{-\eta L(\hat{y}_{t,i}, y_t)}] \right) \\ &= \log \left(\underset{w_{t-1}}{\text{E}} \left[\exp \left(-\eta \left(L(\hat{y}_{t,i}, y_t) - \underset{w_{t-1}}{\text{E}} [L(\hat{y}_{t,i}, y_t)] \right) - \eta \underset{w_{t-1}}{\text{E}} [L(\hat{y}_{t,i}, y_t)] \right) \right] \right) \\ &\leq -\eta \underset{w_{t-1}}{\text{E}} [L(\hat{y}_{t,i}, y_t)] + \frac{\eta^2}{8} \quad (\text{Hoeffding's ineq.}) \\ &\leq -\eta L(\underset{w_{t-1}}{\text{E}} [\hat{y}_{t,i}], y_t) + \frac{\eta^2}{8} \quad (\text{convexity of first arg. of } L) \\ &= -\eta L(\hat{y}_t, y_t) + \frac{\eta^2}{8}.\end{aligned}$$

EW - Proof

- Upper bound: summing up the inequalities yields

$$\Phi_T - \Phi_0 \leq -\eta \sum_{t=1}^T L(\hat{y}_t, y_t) + \frac{\eta^2 T}{8}.$$

- Lower bound:

$$\begin{aligned}\Phi_T - \Phi_0 &= \log \sum_{i=1}^N e^{-\eta L_{T,i}} - \log N \geq \log \max_{i=1}^N e^{-\eta L_{T,i}} - \log N \\ &= -\eta \min_{i=1}^N L_{T,i} - \log N.\end{aligned}$$

- Comparison:

$$\begin{aligned}-\eta \min_{i=1}^N L_{T,i} - \log N &\leq -\eta \sum_{t=1}^T L(\hat{y}_t, y_t) + \frac{\eta^2 T}{8} \\ \Rightarrow \sum_{t=1}^T L(\hat{y}_t, y_t) - \min_{i=1}^N L_{T,i} &\leq \frac{\log N}{\eta} + \frac{\eta T}{8}.\end{aligned}$$

EW - Proof

- **Advantage:** bound on regret per bound is of the form

$$\frac{R_T}{T} = O\left(\sqrt{\frac{\log(N)}{T}}\right).$$

- **Disadvantage:** choice of η requires knowledge of horizon T .

Doubling Trick

- **Idea:** divide time into periods $[2^k, 2^{k+1} - 1]$ of length 2^k with $k = 0, \dots, n$, $T \geq 2^n - 1$, and choose $\eta_k = \sqrt{\frac{8 \log N}{2^k}}$ in each period.
- **Theorem:** with the same assumptions as before, for any T , the following holds:

$$\text{Regret}(T) \leq \frac{\sqrt{2}}{\sqrt{2} - 1} \sqrt{(T/2) \log N} + \sqrt{\log N/2}.$$

Doubling Trick - Proof

- By the previous theorem, for any $I_k = [2^k, 2^{k+1} - 1]$,

$$L_{I_k} - \min_{i=1}^N L_{I_k, i} \leq \sqrt{2^k / 2 \log N}.$$

$$\begin{aligned} \text{Thus, } L_T &= \sum_{k=0}^n L_{I_k} \leq \sum_{k=0}^n \min_{i=1}^N L_{I_k, i} + \sum_{k=0}^n \sqrt{2^k (\log N) / 2} \\ &\leq \min_{i=1}^N L_{T,i} + \sum_{k=0}^n 2^{\frac{k}{2}} \sqrt{(\log N) / 2}. \end{aligned}$$

with

$$\sum_{k=0}^n 2^{\frac{k}{2}} = \frac{\sqrt{2}^{n+1} - 1}{\sqrt{2} - 1} = \frac{2^{(n+1)/2} - 1}{\sqrt{2} - 1} \leq \frac{\sqrt{2}\sqrt{T+1} - 1}{\sqrt{2} - 1} \leq \frac{\sqrt{2}(\sqrt{T} + 1) - 1}{\sqrt{2} - 1} \leq \frac{\sqrt{2}\sqrt{T}}{\sqrt{2} - 1} + 1.$$

Notes

- Doubling trick used in a variety of other contexts and proofs.
- More general method, learning parameter function of time: $\eta_t = \sqrt{(8 \log N)/t}$. Constant factor improvement:

$$\text{Regret}(T) \leq 2\sqrt{(T/2) \log N} + \sqrt{(1/8) \log N}.$$

General Setting

- Adversarial model with action set $X = \{1, \dots, N\}$.
- For $t=1$ to T do
 - player selects distribution p_t over X .
 - adversary selects loss $\mathbf{l}_t = (l_{t,1}, \dots, l_{t,N})$.
 - player receives \mathbf{l}_t .
 - player incurs loss $\sum_{i=1}^N p_{t,i} l_{t,i}$.
- **Objective:** minimize (external) regret

$$R_T = \sum_{t=1}^T \mathbb{E}_{i \sim p_t} [l_{t,i}] - \min_{i=1}^N \sum_{t=1}^T l_{t,i}.$$

Different Setups

- Deterministic vs. randomized.
- Full information vs. bandit setting.
- More general competitor class (e.g., swap regret).
- Oblivious vs. non-oblivious.
- Bounded memory.

Randomized Weighted Majority

(Littlestone and Warmuth, 1988)

RANDOMIZED-WEIGHTED-MAJORITY (N)

```
1  for  $i \leftarrow 1$  to  $N$  do
2       $w_{1,i} \leftarrow 1$ 
3       $p_{1,i} \leftarrow 1/N$ 
4  for  $t \leftarrow 1$  to  $T$  do
5      for  $i \leftarrow 1$  to  $N$  do
6          if ( $l_{t,i} = 1$ ) then
7               $w_{t+1,i} \leftarrow \beta w_{t,i}$ 
8          else  $w_{t+1,i} \leftarrow w_{t,i}$ 
9           $W_{t+1} \leftarrow \sum_{i=1}^N w_{t+1,i}$ 
10         for  $i \leftarrow 1$  to  $N$  do
11              $p_{t+1,i} \leftarrow w_{t+1,i}/W_{t+1}$ 
12 return  $\mathbf{w}_{T+1}$ 
```

RWM - Bound

- **Theorem:** fix $\beta \in (0, 1)$. Then, for any $T \geq 1$, the expected cumulative loss of RWM can be bounded as follows:

$$\mathcal{L}_T \leq \frac{\log N}{1 - \beta} + (2 - \beta)\mathcal{L}_T^{\min}.$$

For $\beta = \max \left\{ \frac{1}{2}, 1 - \sqrt{\frac{\log N}{T}} \right\}$,

$$\mathcal{L}_T \leq \mathcal{L}_T^{\min} + 2\sqrt{T \log N}.$$

RWM - Proof

■ Potential: $W_t = \sum_{i=1}^N w_{t,i}$.

■ Upper bound:

$$\begin{aligned} W_{t+1} &= \sum_{i: l_{t,i}=0} w_{t,i} + \beta \sum_{i: l_{t,i}=1} w_{t,i} = W_t + (\beta - 1) \sum_{i: l_{t,i}=1} w_{t,i} \\ &= W_t + (\beta - 1) W_t \sum_{i: l_{t,i}=1} p_{t,i} \\ &= W_t + (\beta - 1) W_t L_t \\ &= W_t (1 - (1 - \beta) L_t). \end{aligned}$$

Thus, $W_{T+1} = N \prod_{t=1}^T (1 - (1 - \beta) L_t)$.

■ Lower bound: $W_{T+1} \geq \max_{i \in [1, N]} w_{T+1,i} = \beta^{\mathcal{L}_T^{\min}}$.

RWM - Proof

■ Comparison:

$$\beta^{\mathcal{L}_T^{\min}} \leq N \prod_{t=1}^T (1 - (1 - \beta)L_t) \implies \mathcal{L}_T^{\min} \log \beta \leq \log N + \sum_{t=1}^T \log(1 - (1 - \beta)L_t)$$

$$(\forall x < 1, \log(1 - x) \leq -x) \implies \mathcal{L}_T^{\min} \log \beta \leq \log N - (1 - \beta) \sum_{t=1}^T L_t$$

$$\implies \mathcal{L}_T^{\min} \log \beta \leq \log N - (1 - \beta)\mathcal{L}_T$$

$$\implies \mathcal{L}_T \leq \frac{\log N}{1 - \beta} - \frac{\log \beta}{1 - \beta} \mathcal{L}_T^{\min}$$

$$\implies \mathcal{L}_T \leq \frac{\log N}{1 - \beta} - \frac{\log(1 - (1 - \beta))}{1 - \beta} \mathcal{L}_T^{\min}$$

$$(\forall x \in [0, 1/2], -\log(1 - x) \leq x + x^2) \implies \mathcal{L}_T \leq \frac{\log N}{1 - \beta} + (2 - \beta)\mathcal{L}_T^{\min}.$$

■ For the second statement, use

$$\beta^{\mathcal{L}_T^{\min}} \leq \frac{\log N}{1 - \beta} + (2 - \beta)\mathcal{L}_T^{\min} \leq \frac{\log N}{1 - \beta} + (1 - \beta)T + \mathcal{L}_T^{\min}.$$

Lower Bound

- **Theorem:** let $N = 2$. There is a stochastic sequence of losses for which the expected regret of any algorithm verifies $E[R_T] \geq \sqrt{T/8}$.

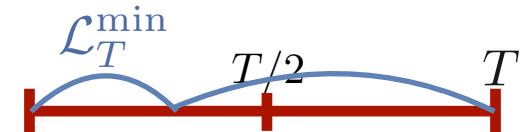
Lower Bound - Proof

- let \mathbf{l}_t take values $\mathbf{l}_{01} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ or $\mathbf{l}_{10} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ with equal probability. Then,

$$\mathbb{E}[\mathcal{L}_T] = \mathbb{E}\left[\sum_{t=1}^T p_t \cdot \mathbf{l}_t\right] = \sum_{t=1}^T p_t \cdot \mathbb{E}[\mathbf{l}_t] = \sum_{t=1}^T \frac{1}{2}p_{t,1} + \frac{1}{2}p_{t,2} = T/2.$$

- Since $\mathcal{L}_{T,1} + \mathcal{L}_{T,2} = T$,

$$\mathcal{L}_T^{\min} = T/2 - |\mathcal{L}_{T,1} - T/2|.$$



- Thus, by the Khintchine-Kahane ineq.,

$$\begin{aligned}\mathbb{E}[R_T] &= \mathbb{E}[\mathcal{L}_T] - \mathbb{E}[\mathcal{L}_T^{\min}] = \mathbb{E}[|\mathcal{L}_{T,1} - T/2|] \\ &= \mathbb{E}\left[\left|\sum_{t=1}^T \frac{1 + \sigma_t}{2} - T/2\right|\right] = \mathbb{E}\left[\left|\sum_{t=1}^T \frac{\sigma_t}{2}\right|\right] \geq \sqrt{T/8}.\end{aligned}$$

Online-to-Batch Conversion

■ Problem:

- sample $((x_1, y_1), \dots, (x_T, y_T)) \in (X \times Y)^T$ drawn i.i.d. according to D .
- loss function L bounded by $M > 0$.
- how do we combine the sequence h_1, \dots, h_{T+1} of hypotheses generated by a regret minimization algorithm to achieve a small generalization error?

Average Generalization

- **Lemma:** for any $\delta > 0$, with probability at least $1 - \delta$, the following holds

$$\frac{1}{T} \sum_{t=1}^T R(h_t) \leq \frac{1}{T} \sum_{t=1}^T L(h_t(x_t), y_t) + M \sqrt{\frac{2 \log \frac{1}{\delta}}{T}}.$$

- **Proof:** for any $t \in [1, T]$, let $V_t = R(h_t) - L(h_t(x_t), y_t)$.
 - Then,
$$\mathbb{E}[V_t | x_{1:t-1}] = R(h_t) - \mathbb{E}[L(h_t(x_t), y_t) | x_{1:t-1}] = R(h_t) - R(h_t) = 0.$$
 - By Azuma's inequality,

$$\Pr\left[\frac{1}{T} \sum_{i=1}^T V_i \geq \epsilon\right] \leq \exp(-2T\epsilon^2/(2M)^2)).$$

Online-to-Batch Guarantee

- **Theorem:** for any $\delta > 0$, with probability at least $1 - \delta$, the following holds for a convex loss:

$$R\left(\frac{1}{T} \sum_{i=1}^T h_i\right) \leq \frac{1}{T} \sum_{i=1}^T L(h_i(x_i), y_i) + M \sqrt{\frac{2 \log \frac{1}{\delta}}{T}}$$

$$R\left(\frac{1}{T} \sum_{i=1}^T h_i\right) \leq \inf_{h \in H} R(h) + \frac{R_T}{T} + 2M \sqrt{\frac{2 \log \frac{2}{\delta}}{T}}.$$

Proof

- Convexity: $L\left(\frac{1}{T} \sum_{i=1}^T h_i(x), y\right) \leq \frac{1}{T} \sum_{i=1}^T L(h_i(x), y).$
Thus,
$$R\left(\frac{1}{T} \sum_{i=1}^T h_i\right) \leq \frac{1}{T} \sum_{i=1}^T R(h_i).$$
- By definition of the regret, with probability at least $1 - \delta/2$,

$$\begin{aligned} R\left(\frac{1}{T} \sum_{i=1}^T h_i\right) &\leq \frac{1}{T} \sum_{i=1}^T L(h_i(x_i), y_i) + M \sqrt{\frac{2 \log \frac{2}{\delta}}{T}} \\ &\leq \inf_{h \in H} \frac{1}{T} \sum_{i=1}^T L(h(x_i), y_i) + \frac{R_T}{T} + M \sqrt{\frac{2 \log \frac{2}{\delta}}{T}}. \end{aligned}$$

Proof

- Assume that the infimum is reached at h^* . By Hoeffding's inequality, with probability at least $1 - \delta/2$,

$$\frac{1}{T} \sum_{i=1}^T L(h^*(x_i), y_i) \leq R(h^*) + M \sqrt{\frac{2 \log \frac{2}{\delta}}{T}}.$$

- Thus, with probability at least $1 - \delta$,

$$\begin{aligned} R\left(\frac{1}{T} \sum_{i=1}^T h_i\right) &\leq \frac{1}{T} \sum_{i=1}^T L(h^*(x_i), y_i) + \frac{R_T}{T} + M \sqrt{\frac{2 \log \frac{2}{\delta}}{T}} \\ &\leq R(h^*) + M \sqrt{\frac{2 \log \frac{2}{\delta}}{T}} + \frac{R_T}{T} + M \sqrt{\frac{2 \log \frac{2}{\delta}}{T}} \\ &= R(h^*) + \frac{R_T}{T} + 2M \sqrt{\frac{2 \log \frac{2}{\delta}}{T}}. \end{aligned}$$

References

- Nicolò Cesa-Bianchi, Alex Conconi, Claudio Gentile: On the Generalization Ability of On-Line Learning Algorithms. *IEEE Transactions on Information Theory* 50(9): 2050-2057. 2004.
- Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.
- Yoav Freund and Robert Schapire. Large margin classification using the perceptron algorithm. In *Proceedings of COLT 1998*. ACM Press, 1998.
- Adam T. Kalai, Santosh Vempala. Efficient algorithms for online decision problems. *J. Comput. Syst. Sci.* 71(3): 291-307. 2005.
- Nick Littlestone. From On-Line to Batch Learning. *COLT 1989*: 269-284.
- Nick Littlestone. "Learning Quickly When Irrelevant Attributes Abound: A New Linear-threshold Algorithm" *Machine Learning* 285-318(2). 1988.

References

- Nick Littlestone, Manfred K. Warmuth: The Weighted Majority Algorithm. *FOCS* 1989: 256-261.
- Tom Mitchell. *Machine Learning*, McGraw Hill, 1997.
- Novikoff, A. B. (1962). On convergence proofs on perceptrons. *Symposium on the Mathematical Theory of Automata*, 12, 615-622. Polytechnic Institute of Brooklyn.