INTRODUCCIÓN AL APRENDIZAJE ON LINE

PROBABILIDAD 2023 - RICARDO FRAIMAN

1. Introducción

Consideremos el problema de predecir una secuencia desconocida

$$\{y_1, y_2, \ldots\},$$

de bits $y_i \in \{0,1\}$. En cada tiempo t el pronosticador primero hace su conjetura $\widehat{p}_t \in \{0,1\}$. Luego el bit verdadero y_t es revelado y el pronosticador observa si su predicción fue correcta. Para calcular $\widehat{p}_t$ el pronosticador atiende al consejo de N expertos. El consejo toma la forma de un vector binario $(f_{1t}, \ldots, f_{Nt})$, donde $f_{it} \in \{0,1\}$ es la predicción del experto i para el próximo bit y_t.

El objetivo es acotar el número de instantes t en que $\widehat{p}_t \neq y_t$, o sea acotar el número de errores que realiza el pronosticador.

Comenzaremos con un caso aún mas simple, en que supondremos que sabemos que en la secuencia $(f_{1t}, \ldots, f_{Nt})$ hay uno correspondiente algún experto i que no se equivoca jamás. O sea, sabemos que $f_{it} = y_t$ para cierto i para todo t pero no sabemos cual es el i para el cual se verifica.

Usando esta información, es fácil ver que el pronosticador puede elaborar una estrategia que hace a lo sumo $\lfloor log_2(N) \rfloor$ errores en la secuencia.

En efecto, supongamos que la estrategia del pronosticador comienza asignando un peso $w_j = 1$ a cada experto, $j = 1, \ldots, N$. En cada paso t, el pronosticador predice $\widehat{p}_t = 1$ si y solo si el número de expertos con peso $w_j = 1$ que proponen $f_{jt} = 1$ es mayor que el de aquellos con $w_j = 1$ que proponen $f_{jt} = 0$.

Luego que y_t es revelado, si $\widehat{p}_t \neq y_t$ el pronosticador reasigna pesos $w_k \leftarrow 0$ para todos los expertos k cuyos $f_{kt} \neq y_t$. O sea, el pronosticador va observando los expertos que han cometido un error y se queda solo con aquellos que no lo han hecho para predecir en el siguiente paso.

Sea W_m la suma de los pesos de todos los expertos luego que el pronosticador ha cometido su m-esimo error. Esto implica que $W_m \leq W_{m-1}/2$, pues aquellos expertos que fueron considerados incorrectos por primera vez tienen sus pesos llevados a cero por el pronosticador y son al menos la mitad. Como la desigualdad vale para todo $m \geq 1$, tenemos que $W_m \leq W_0/2^m$.

Como sabemos que uno de los expertos i nunca se equivoca, entonces $w_i = 1$ y por tanto $W_m \geq 1$. Finalmente como $W_0 = N$, tenemos que $1 \leq N/2^m$, y despejando m tenemos el resultado.

2. Un caso más general

Consideremos ahora el caso general, en que el pronosticador no tiene ninguna información preliminar sobre la cantidad de errores que los expertos cometerán en la secuencia. El objetivo ahora es relacionar el número de errores que realiza el pronosticador, con el número de errores que comete el mejor de los expertos, independientemente de cual sea la secuencia a predecir.

En este caso tendremos que modificar la estrategia, pues llevar a cero los pesos de quienes se equivocan solo hace sentido cuando estamos seguros que uno de ellos no se equivoca jamás.

Sin esta certeza, una estrategia mas segura es realizar la asignación $w_k \leftarrow \beta w_k$ cada vez que el experto k cometa un error, donde $0 < \beta < 1$ es un parámetro libre. Esta es la única modificación que realizaremos a la estrategia del pronosticador utilizada antes.

Mas precisamente, el nuevo pronóstico compara la suma de pesos de los expertos que recomiendan predecir 1 con la de aquellos que recomiendan predecir 0, y predice de acuerdo a cual suma es mayor.

Como antes, cuando el pronosticador comete su m-esimo error, la suma de pesos de los expertos que se equivocaron es al menos $W_{m-1}/2$. Los pesos de dichos expertos se multiplican por β y los del resto de los expertos, que a lo sumo es $W_{m-1}/2$ permanecen incambiados. Tenemos entonces que $W_m \leq W_{m-1}/2 + \beta W_{m-1}/2$. Como esto vale para todo $m \geq 1$, obtenemos que $W_m \leq W_0(1+\beta)^m/2^m$.

Sea k el experto que ha cometido menos errores cuando el pronosticador comete su m-ésimo error. Denotemos por m^* la cantidad mínima de errores cometidos por dicho experto. Entonces el peso actual de este experto es $w_k = \beta^{m^*}$, y por tanto tenemos que $W_m \geq \beta^{m^*}$. Esto nos provee la desigualdad,

$$\beta^{m^*} \leq W_0(1+\beta)^m/2^m.$$

Finalmente, usando además que $W_0 = N$, obtenemos que

$$m \leq \left\lfloor \frac{log_2(N) + m^* log_2(1/\beta)}{log_2(2/(1+\beta))} \right\rfloor .$$

Para cualquier valor fijo de β, esta desigualdad establece una dependencia lineal entre los errores cometidos por el pronosticador, luego de cualquier cantidad de pronósticos, y los errores cometidos por el mejor experto luego de la misma cantidad de pronósticos.

3. Un marco un poco mas general

Se repite un juego entre el pronosticador y el ambiente (the forcarster and the environment). Para cada tiempo t

- el pronosticador elige una acción $I_t \in \{1, \ldots, N\}$ (a menudo las acciones se llaman expertos)
- el ambiente elige una pérdida $l_t(1), l_t(2), \ldots, l_t(N) \in [0, 1]$.
- el pronosticador sufre una pérdida $l_t(I_t)$

El objetivo es buscar una estrategia que minimize el arrepentimiento (regret)

$$R_n = \left(\sum_{t=1}^{n} l_t(I_t) - \min_{i \leq N} \sum_{t=1}^{n} l_t(i) \right), \tag{1}$$

o sea que la estrategia se parezca a la del mejor de los expertos.

Es posible lograr que $R_n/n \to 0$ para toda asignación de pérdidas por el ambiente?

Example 3.1. Tomemos $N = 2$ y definimos para todo t

$$l_t(1) = \begin{cases} 0 & \text{si } I_t = 2 \\ 1 & \text{si } I_t = 1 \end{cases} \tag{2}$$

$l_t(2) = 1 - l_t(1)$. Entonces $\sum_{t=1}^{n} l_t(I_t) = n$ y $\min_{i=1,2} \sum_{t=1}^{n} l_t(i) \leq n/2$, por tanto $\frac{1}{n} R_n \geq 1/2$.

4. Clave para la solución: predicción aleatorizada.

En cada instante t el pronosticador elige una distribución de probabilidad $p_{t-1} = (p_{1,t-1}, \ldots, p_{N,t-1})$ y elige la acción i con probabilidad $p_{i,t-1}$.

Supondremos además el caso mas simple en que todas las pérdidas $l_s(i)$, $i = 1, \ldots, N$, $s < t$ son observadas.

Theorem 4.1. *(Hannan (1957) y Blackell (1956)) El pronosticador tiene una estrategia tal que*

$$\frac{1}{n} \left(\sum_{t=1}^{n} l_t(I_t) - \min_{i \leq N} \sum_{t=1}^{n} l_t(i) \right) \to 0 \ \textit{a.s.},$$

para todas las estrategias del ambiente.

En este caso se puede probar que una estrategia que asigne mayor probabilidad a las acciones que tienen mejor performance (Vouk (1990). Littlestone-Warmuth (1989)) de la forma

$$p_{i,t-1} = \frac{exp\left(-\theta \sum_{s=1}^{t-1} l_s(i)\right)}{\sum_{k=1}^{N} exp\left(-\theta \sum_{s=1}^{t-1} l_s(k)\right)},$$

para $i = 1, \dots N$ donde $\theta > 0$, verifica que

$$\frac{1}{n}\left(\sum_{i=1}^{n} l_t(p_{t-1}) - \min_{i \leq N} \sum_{t=1}^{n} l_t(i)\right) = \sqrt{\frac{lnN}{2n}}, \tag{3}$$

con $\theta = \sqrt{8ln(N)/n}$.

La demostración requiere de pre-requisitos sobre martingalas que exceden este curso.

Algunas otras propuestas.

Hannan (1957). Follow de perturbed leader

$$I_t = argmin_{1 \leq i \leq N} \sum_{s=1}^{t-1} l_s(i) + Z_{it},$$

donde Z_{it} son v.a. (ruidos).

- Si Z_{it} son iid $U(0, \sqrt{nN})$ entonces
$$\frac{1}{n} R_n \leq 2\sqrt{\frac{N}{n}} + O_P(n^{-1/2}).$$
- Si Z_{it} son iid con densidad $\frac{\theta}{2} e^{-\theta|z|}$, entonces para $\theta \sim \sqrt{ln(N)/n}$
$$\frac{1}{n} R_n \leq C\sqrt{\frac{N}{n}} + O_P(n^{-1/2}).$$

(Kalai-Vempala (2003)