

## 2

### *Previsão com aconselhamento especializado*

O modelo de previsão com aconselhamento especializado, apresentado neste capítulo, fornece as bases para a teoria de previsão de sequências individuais que desenvolvemos no restante do livro.

A previsão com aconselhamento especializado é baseada no seguinte protocolo para decisões sequenciais: o tomador de decisão é um previsor cujo objetivo é prever uma sequência desconhecida de elementos de um *espaço de resultados*  $Y$ . As previsões do previsor  $p_1, p_2, \dots, y_1, y_2, \dots$  pertencem a um *espaço de decisão*  $D$ , assumimos ser um subconjunto convexo de um espaço vetorial. Em alguns casos especiais, tomamos  $D = Y$ , mas em geral  $D$  pode ser diferente de  $Y$ .

O meteorologista calcula suas previsões de forma sequencial, e seu desempenho preditivo é comparado ao de um conjunto de meteorologistas de referência que chamamos de *especialistas*.

Mais precisamente, a cada instante  $t$  o previsor tem acesso ao conjunto  $f_{E,t} : E \rightarrow D$  de previsões de especialistas  $f_{E,t} \in D$ , onde  $E$  é um conjunto fixo de índices para os especialistas. Com base nas previsões dos especialistas, o previsor calcula seu próprio palpite  $p_t$  para o próximo resultado  $y_t$ .

Depois que  $p_t$  é calculado, o resultado verdadeiro  $y_t$  é revelado.

As previsões dos analistas e especialistas são pontuadas usando uma *função de perda* não negativa  $l : D \times Y \rightarrow \mathbb{R}$ .

Este protocolo de previsão pode ser naturalmente visto como o seguinte jogo repetido entre o "previsor", que faz suposições  $p_t$ , e o "ambiente", que escolhe o conselho especializado  $f_{E,t} : E \rightarrow D$  e define os resultados verdadeiros  $y_t$ .

#### PREVISÃO COM ACONSELHAMENTO DE ESPECIALISTAS

**Parâmetros:** espaço de decisão  $D$ , espaço de resultado  $Y$ , índices de função de perda  $l$ , conjunto  $E$  de especialista

Para cada rodada  $t = 1, 2, \dots$

- (1) o ambiente escolhe o próximo resultado  $y_t$  e o conselho do especialista  $f_{E,t} : E \rightarrow D$ ; o conselho do especialista é revelado ao previsor;
- (2) o previsor escolhe a previsão  $p_t \in D$ ; (3) o ambiente revela o próximo resultado  $y_t \in Y$ ; (4) o previsor incorre em perda  $l(p_t, y_t)$  e cada especialista  $E$  incorre em perda  $l(f_{E,t}, y_t)$ .

O objetivo do previsor é manter o menor possível o *arrependimento cumulativo* (ou simplesmente *arrependimento*) com relação a cada especialista. Essa quantidade é definida, para o especialista  $E$ , pela soma

$$RE,n = \sum_{t=1}^n (p_t(y_t) - f_E(t, y_t)) = L_n - L_{E,n}$$

onde usamos  $L_n = \sum_{t=1}^n p_t(y_t)$  para denotar a perda cumulativa do previsor e  $L_{E,n} = \sum_{t=1}^n f_E(t, y_t)$  para denotar a perda cumulativa do especialista  $E$ . Portanto,  $RE,n$  é a diferença entre a perda total do previsor e a do especialista  $E$  após  $n$  rodadas de previsão. Também definamos o arrependimento instantâneo em relação ao especialista  $E$  no tempo  $t$  por  $r_{E,t} = p_t(y_t) - f_E(t, y_t)$ , de modo que  $RE,n = \sum_{t=1}^n r_{E,t}$ . Pode-se pensar em  $r_{E,t}$  como o arrependimento do previsor se não ter ouvido o conselho do especialista  $E$  logo após o resultado  $y_t$  ter sido revelado.

Ao longo do resto deste capítulo, assumimos que o número de especialistas é finito,  $E = \{1, 2, \dots, N\}$ , e usamos o índice  $i = 1, \dots, N$  para se referir a um especialista. O objetivo do previsor deve prever de modo que o arrependimento seja o menor possível para todas as sequências de resultados. Por exemplo, o previsor pode querer ter um arrependimento por rodada que desaparece, ou seja, para alcançar

$$\max_{E=1,\dots,N} RE,n = o(n) \text{ ou, equivalentemente, } \frac{1}{n} L_n - \min_{E=1,\dots,N} L_{E,n} \rightarrow 0,$$

onde a convergência é uniforme sobre a escolha da sequência de resultados e a escolha do conselho do especialista. Na próxima seção mostramos que esta meta ambiciosa pode ser alcançada por um simples meteorologista em condições amenas.

O restante do capítulo está estruturado da seguinte forma. Na Seção 2.1, apresentamos os importantes classes de previsores de média ponderada, descrevemos uma subclasse de previsores baseados em potencial, e analisamos dois casos especiais importantes: o previsor de média ponderada polinomial e o previsor de média ponderada exponencialmente. Este último previsor é bastante central em nossa teoria, e as quatro seções seguintes estão todas preocupadas com várias questões relacionadas a ele: a Seção 2.2 mostra certas propriedades de otimalidade, a Seção 2.3 aborda o problema de ajuste dinâmico do parâmetro do potencial, a Seção 2.4 investiga o problema de obter limites de arrependimento melhorados quando a perda do melhor especialista é pequeno, e a Seção 2.5 investiga o caso especial de funções de perda diferenciáveis. Começando com a Seção 2.6, descobrimos as vantagens de redimensionar a função de perda. Este truque simples nos permite derivar limites de desempenho novos e ainda mais nítidos. Na Seção 2.7 introduzimos e analisamos um previsor de média ponderada para perdas redimensionadas que, ao contrário dos anteriores, não se baseia na noção de potencial. Na Seção 2.8 retornamos ao previsor de média ponderada exponencialmente e derivamos limites de arrependimento aprimorados com base em redimensionando a função de perda. As seções 2.9 e 2.10 abordam algumas questões gerais no problema de previsão com aconselhamento especializado, incluindo a definição de valores minimax. Finalmente, na Seção 2.11 discutimos uma variante da noção de arrependimento onde os fatores de desconto são introduzidos.

## 2.1 Previsão de média ponderada

Uma estratégia de previsão natural neste quadro é baseada no cálculo de uma *média ponderada* das previsões dos especialistas. Ou seja, o previsior prevê no tempo  $t$  de acordo com

$$p_t = \frac{\sum_{i=1}^N w_{i,t} f_{i,t}}{\sum_{j=1}^N w_{j,t}},$$

onde  $w_1, t, \dots, w_N, t \geq 0$  são os pesos atribuídos aos especialistas no tempo  $t$ . Observe que  $p_t \in D$ , uma vez que é uma combinação convexa do conselho especializado  $f_1, t, \dots, f_N, t \in D$  e  $D$  é convexo por nossas suposições. Como nosso objetivo é minimizar o arrependimento, é razoável escolher os pesos de acordo com o arrependimento até o tempo  $t-1$ . Se  $R_{i,t-1}$  for grande, então nós atribuir um grande peso  $w_{i,t}$  ao especialista  $i$ , e vice-versa. Como  $R_{i,t-1} = L_{i,t-1} - L_{i,t}$ , isso resulta em ponderar mais aqueles especialistas  $i$  cuja perda cumulativa  $L_{i,t}$  é pequena. Portanto, nós veja o peso como uma função crescente arbitrária do arrependimento do especialista. Por razões que ficará claro em breve, achamos conveniente escrever esta função como a derivada de uma função não negativa, convexa e crescente  $\tilde{y} : \mathbb{R} \rightarrow \mathbb{R}$ . Escrevemos  $\tilde{y}$  para denotar isso derivada. O previsior usa  $\tilde{y}$  para determinar o peso  $w_{i,t} = \tilde{y}(R_{i,t-1})$  atribuído a o  $i$ -ésimo especialista. Portanto, a previsão  $p_t$  no tempo  $t$  do previsior médio ponderado é definido por

$$p_t = \frac{\sum_{i=1}^N \tilde{y}(R_{i,t-1}) f_{i,t}}{\sum_{j=1}^N \tilde{y}(R_{j,t-1})} \quad (\text{previsior de média ponderada}).$$

Observe que este é um previsior legítimo, pois o  $p_t$  é calculado com base no conselho dos especialistas no momento  $t$  e os arrependimentos cumulativos até o momento  $t-1$ .

Começamos a análise dos meteorologistas de média ponderada por uma simples observação técnica.

**Lema 2.1.** Se a função de perda for convexa em seu primeiro argumento, então

$$\tilde{y}(R_{i,t-1}) - \tilde{y}(R_{j,t-1}) \leq \sum_{s=1}^{t-1} (f_{i,s} - f_{j,s}) \quad \text{e} \quad \tilde{y}(R_{i,t-1}) \geq 0.$$

**Prova.** Usando a desigualdade de Jensen, para todo  $y \in Y$ ,

$$(p_t, e) = \frac{\sum_{i=1}^N \tilde{y}(R_{i,t-1}) f_{i,t}}{\sum_{j=1}^N \tilde{y}(R_{j,t-1})} \cdot y \leq \frac{\sum_{i=1}^N \tilde{y}(R_{i,t-1}) (f_{i,t}, y)}{\sum_{j=1}^N \tilde{y}(R_{j,t-1})}.$$

Reorganizando, obtemos a afirmação. ■

A simples observação do lema acima nos permite interpretar a média ponderada previsior de uma forma interessante. Para fazer isso, introduza o vetor de *arrependimento instantâneo*

$$r_t = (r_{1,t}, \dots, r_{N,t}) \in \mathbb{R}^N$$

e o vetor de arrependimento correspondente  $R_n = \sum_{t=1}^n r_t$ . É conveniente introduzir também uma função potencial :  $RN \rightarrow \mathbb{R}$  da forma

$$(u) = \sum_{i=1}^N \tilde{y}(ui) \quad (\text{função potencial}),$$

onde  $\tilde{y} : \mathbb{R} \rightarrow \mathbb{R}$  é qualquer função não negativa, crescente e duas vezes diferenciável, e  $\tilde{y} : \mathbb{R} \rightarrow \mathbb{R}$  é qualquer função auxiliar não negativa, estritamente crescente, côncava e duas vezes diferenciável função.

Usando a noção de função potencial, podemos dar a seguinte definição equivalente do predictor médio ponderado

$$p_t = \frac{\sum_{i=1}^N \tilde{y}(R_t \tilde{y}^{-1}(i))}{\sum_{j=1}^N \tilde{y}(R_t \tilde{y}^{-1}(j))}$$

onde  $\tilde{y}(R_t \tilde{y}^{-1}(i)) = \tilde{y}(R_t \tilde{y}^{-1}(i)) / \tilde{y}(R_t \tilde{y}^{-1}(i))$ . Dizemos que um predictor definido como acima é *baseado sobre o potencial*. Embora a definição do predictor de média ponderada seja independente da escolha de  $\tilde{y}$  (as derivadas  $\tilde{y}$  cancelam na definição de  $p_t$  acima), a prova do resultado principal deste capítulo, Teorema 2.1, revela que  $\tilde{y}$  desempenha um papel importante na análise. Observamos que a convexidade de  $\tilde{y}$  não é necessária para provar o Teorema 2.1, e isso é a razão pela qual a convexidade não é mencionada na definição acima de função potencial. Por outro lado, todos os predictors deste livro que se baseiam em funções potenciais e têm um arrependimento por rodada que desaparece e são construídos usando um  $\tilde{y}$  convexo (veja também Exercício 2.2).

A afirmação do Lema 2.1 é equivalente a

$$r_t \cdot \tilde{y}(R_t \tilde{y}^{-1}(i)) \leq 0 \quad \text{sup}_{y \in Y} \quad (\text{Condição de Blackwell}).$$

A notação  $u \cdot v$  representa o produto interno de dois vetores definidos por  $u \cdot v = u_1 v_1 + \dots + u_N v_N$ . Chamamos a desigualdade acima de *condição de Blackwell* por causa de sua semelhança com uma propriedade chave usada na prova da celebrada acessibilidade de Blackwell teorema. O teorema e sua conexão com a desigualdade acima são explorados nas Seções 7.7 e 7.8. A Figura 2.1 mostra um exemplo de uma previsão que satisfaz o Blackwell doença.

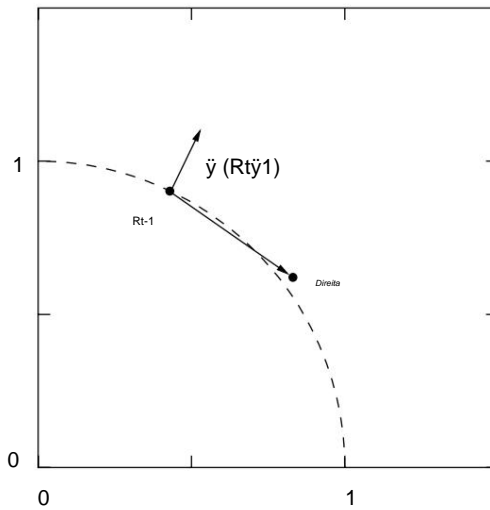
A condição de Blackwell mostra que a função desempenha um papel vagamente semelhante ao potencial em um sistema dinâmico: o predictor de média ponderada, forçando o arrependimento vetor para apontar para longe do gradiente de independentemente do resultado  $y_t$ , tende a manter o ponto  $R_t$  próximo ao mínimo de Esta propriedade, de fato, sugere uma análise simples porque os incrementos da função potencial podem agora ser facilmente limitados pela função de Taylor teorema. O papel da função  $\tilde{y}$  é simplesmente obter melhores limites com este argumento.

O próximo teorema se aplica a qualquer predictor que satisfaça a condição de Blackwell (e, portanto, não apenas para os analistas de média ponderada). No entanto, isso implicará vários limites interessantes para diferentes versões do predictor de média ponderada.

**Teorema 2.1.** *Suponha que um predictor satisfaça a condição de Blackwell para um potencial*

$$(u) = \sum_{i=1}^N \tilde{y}(ui). \text{ Então, para todo } n = 1, 2, \dots,$$

$$(R_n) \leq (0) + \frac{1}{2} \sum_{t=1}^n C(r_t),$$



**Figura 2.1.** Uma ilustração da condição de Blackwell com  $N = 2$ . A linha tracejada mostra os pontos no espaço de arrependimento com potencial igual a 1. A previsão no tempo  $t$  mudou o potencial de  $(R_t \hat{y}) = 1$  para  $(R_t) = (R_t \hat{y} + r_t)$ . Embora  $(R_t) > (R_t \hat{y})$ , o produto interno entre  $r_t$  e o gradiente  $\hat{y}(R_t \hat{y})$  seja negativo e, portanto, a condição de Blackwell se mantém.

onde

$$C(r_t) = \sup_{\substack{\text{você } \hat{y} \in RN \\ EU=1}} \hat{y} \cdot \hat{y}(u_i) - \sum_{\substack{EU=1 \\ \text{isto}}} \hat{y}(u_i) r_i^2$$

**Prova.** Estimamos  $(R_t)$  em termos de  $(R_t \hat{y})$  usando o teorema de Taylor. Assim, obtemos

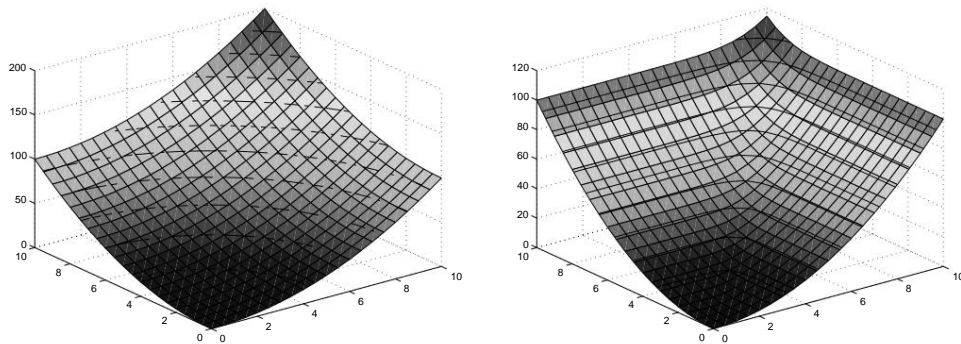
$$\begin{aligned} (R_t) &= (R_t \hat{y} + r_t) \\ &= (R_t \hat{y}) + \hat{y}(R_t \hat{y}) \cdot r_t + \frac{1}{2} \sum_{\substack{EU=1 \\ j=1}}^{N-1} \frac{\hat{y}_j^2}{\hat{y}_{ui} \hat{y}_{uj}} r_i r_j, t \\ &\quad (\text{onde } \hat{y} \text{ é algum vetor em } RN) \\ &= \hat{y}(R_t \hat{y}) + 2 \sum_{\substack{EU=1 \\ j=1}}^{N-1} \frac{\hat{y}_j^2}{\hat{y}_{ui} \hat{y}_{uj}} r_i r_j, t \end{aligned}$$

onde a desigualdade segue pela condição de Blackwell. Agora, um cálculo direto mostra que

$$\begin{aligned} &\sum_{\substack{EU=1 \\ j=1}}^{N-1} \frac{\hat{y}_j^2}{\hat{y}_{ui} \hat{y}_{uj}} r_i r_j, t \\ &= \sum_{\substack{EU=1 \\ j=1}}^{N-1} \hat{y}(\hat{y}_i) \hat{y}(\hat{y}_j) r_i r_j, t \\ &\quad + \sum_{\substack{EU=1 \\ j=1}}^{N-1} \hat{y}(\hat{y}_i) \hat{y}(\hat{y}_j) r_i^2 \end{aligned}$$

**Corolário 2.1.** *Suponha que a função de perda seja convexa em seu primeiro argumento e que ela assuma valores em  $[0, 1]$ . Então, para qualquer sequência  $y_1, y_2, \dots \ddot{Y}$  de resultados e para qualquer  $n \ddot{Y} 1$ , a*

**Observação 2.1.** Definimos o potencial polinomial como  $\mathbf{p}(\mathbf{u}) = \mathbf{u} +$  corresponde <sup>2</sup> que corresponde , a tomar  $\tilde{y}(x) = x/2p$ . Lembre-se de que  $\tilde{y}$  não tem nenhuma influência na previsão, ele só tem um papel na análise. A forma particular analisada aqui é escolhida por conveniência, mas há outras possibilidades que levam a resultados semelhantes. Por exemplo, pode-se argumentar que é mais natural tomar  $\tilde{y}(x) = x/1/p$ , o que leva à função potencial  $(\mathbf{u}) = \mathbf{u} + p$ .



**Figura 2.2.** Gráficos da função potencial polinomial  $p(u)$  para  $N = 2$  especialistas com expoentes  $p = 2$  e  $p = 10$ .

Deixamos como exercício a tarefa de descobrir um limite semelhante ao do Corolário 2.1 com base nessa escolha.

### Previsor de média ponderada exponencialmente

Nosso segundo exemplo principal é o *previsor de média ponderada exponencialmente* com base no potencial

$$\tilde{y}(u) = \sum_{i=1}^N \frac{e^{\tilde{y} u_i}}{e^{\tilde{y} u_i} + 1} \quad (\text{potencial exponencial}),$$

onde  $\tilde{y}$  é um parâmetro positivo. Neste caso, os pesos atribuídos aos especialistas são da forma

$$w_{i,t\tilde{y}} = \tilde{y} (R_{i,t\tilde{y}})^i = \frac{e^{\tilde{y} R_{i,t\tilde{y}}}}{\sum_{j=1}^N e^{\tilde{y} R_{j,t\tilde{y}}}},$$

e o previsor de média ponderada simplifica para

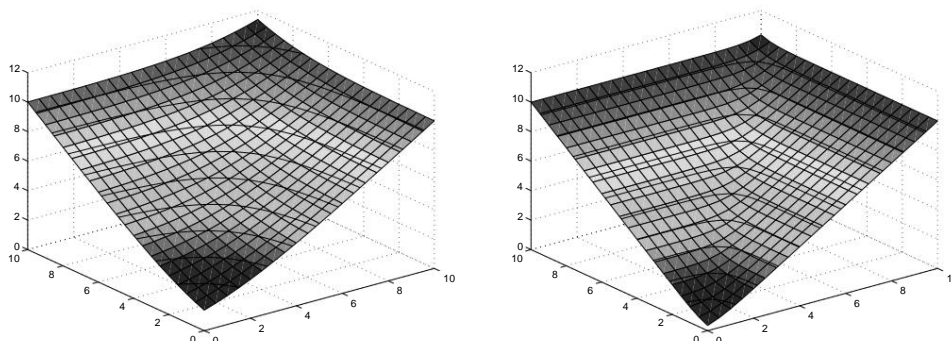
$$p_t = \frac{\sum_{i=1}^N \exp \tilde{y} L_{i,t\tilde{y}} \tilde{y} L_{i,t\tilde{y}} f_{i,t}}{\sum_{j=1}^N \exp \tilde{y} L_{j,t\tilde{y}} \tilde{y} L_{j,t\tilde{y}}} = \frac{\sum_{i=1}^N e^{\tilde{y} L_{i,t\tilde{y}} \tilde{y} L_{i,t\tilde{y}} f_{i,t}}}{\sum_{j=1}^N e^{\tilde{y} L_{j,t\tilde{y}} \tilde{y} L_{j,t\tilde{y}}}}.$$

A beleza do previsor de média ponderada exponencialmente é que ele depende apenas do desempenho passado dos especialistas, enquanto as previsões feitas usando outros potenciais gerais dependem das previsões passadas  $p_{s,s} < t$ , também. Além disso, os pesos que o previsor atribui aos especialistas são computáveis de uma forma incremental simples: deixe  $w_{1,t\tilde{y}}, \dots, w_{N,t\tilde{y}}$  ser os pesos usados na rodada  $t$  para calcular a previsão  $p_t$  = pode-se verificar facilmente,  $\sum_{i=1}^N w_{i,t\tilde{y}} f_{i,t} = p_t$ . Então, como

$$com_t = \frac{\sum_{i=1}^N w_{i,t\tilde{y}} e^{\tilde{y} L_{i,t\tilde{y}} \tilde{y} L_{i,t\tilde{y}} f_{i,t}}}{\sum_{j=1}^N w_{j,t\tilde{y}} e^{\tilde{y} L_{j,t\tilde{y}} \tilde{y} L_{j,t\tilde{y}} f_{j,t}}} \cdot (f_{i,t}, y_t) \cdot w_{i,t\tilde{y}}.$$

Uma aplicação simples do Teorema 2.1 revela o seguinte limite de desempenho para o previsor de média ponderada exponencialmente.





**Figura 2.3.** Gráficos da função potencial exponencial  $\tilde{y}(u)$  para  $N = 2$  especialistas com  $\tilde{\gamma} = 0,5$  e  $\tilde{\gamma} = 2$ .

**Corolário 2.2.** Suponha que a função de perda seja convexa em seu primeiro argumento e que ela,  $y \in Y$ , o valores em  $[0, 1]$ . Para qualquer  $n$  e  $\tilde{\gamma} > 0$ , e para todos  $y_1, \dots$  o previsor de média arrendimento de assume ponderada exponencialmente satisfaz

$$L_n \tilde{\gamma} \min_{i=1, \dots, N} \quad L_{i,n} \tilde{\gamma} + \frac{\text{em } N}{\tilde{\gamma}} \quad \frac{\text{não}}{2}.$$

Otimizar o limite superior sugere a escolha  $\tilde{\gamma} = \sqrt{2 \ln N / n}$ . Neste caso, o limite superior se torna  $\sqrt{2 \ln N}$ , que é ligeiramente melhor do que o melhor limite que obtivemos usando  $\tilde{y}(x) = x$  com  $p = 2 \ln N$ . Na próxima seção, melhoramos o limite do Corolário 2.2 por uma análise direta. A desvantagem da ponderação exponencial é que o ajuste ótimo do parâmetro  $\tilde{\gamma}$  requer conhecimento do horizonte  $n$  com antecedência. Nas próximas duas seções, descrevemos versões do previsor de média ponderada exponencial que não sofrem dessa desvantagem.

**Prova do Corolário 2.2.** Aplique o Teorema 2.1 usando o potencial exponencial. Então  $\tilde{y}(x) = e^{\tilde{\gamma}x}$ ,  $\tilde{y}(x) = (1/\tilde{\gamma}) \ln x$ , e

$$\tilde{y} \quad \tilde{y}(u_i) \quad \tilde{y}(u_i) \tilde{\gamma} \tilde{\gamma} \max_{i=1, \dots, N} \quad \frac{2}{e u, t \tilde{\gamma} \tilde{\gamma}}.$$

Usando  $\tilde{y}(0) = (\ln N)/\tilde{\gamma}$ , o Teorema 2.1 implica que

$$\max_{i=1, \dots, N} R_{i,n} \tilde{\gamma} \tilde{y}(R_n) \tilde{\gamma} + 2 \quad \frac{\text{em } N \tilde{\gamma}}{\tilde{\gamma}} \quad \frac{\text{não}}{2}.$$

conforme desejado. ■

## 2.2 Um limite ideal

O objetivo desta seção é mostrar que, mesmo para funções de perda convexas gerais, o limite do Corolário 2.2 pode ser melhorado para o previsor de média ponderada exponencialmente.

O resultado a seguir melhora o Corolário 2.2 por um fator constante. Na Seção 3.7 vemos que o limite obtido aqui não pode ser melhorado mais.

**Teorema 2.2.** Suponha que a função de perda seja convexa em seu primeiro argumento e que ela assume valores em  $[0, 1]$ . Para qualquer  $n$  e  $\gamma > 0$ , e para todos  $y_1, \dots, y_n \in Y$ , o arrependimento do o preditor de média ponderada exponencial satisfaz

$$L_n \leq \min_{u=1, \dots, N} L_{i,n} \leq \frac{\ln N}{\gamma} + \frac{1}{8}.$$

Em particular, com  $\gamma = \frac{1}{8 \ln N/n}$ , o limite superior se torna  $\frac{1}{2} \ln N$ .

A prova é semelhante, em espírito, à do Corolário 2.2, mas agora, em vez de limitar o crescimento de  $(1/\gamma) \ln$  limitamos as quantidades relacionadas  $(1/\gamma) \ln(W_t/W_{t-1})$ , onde

$$W_t = \sum_{i=1}^N e^{-\gamma L_{i,t}}$$

para  $t \geq 1$ , e  $W_0 = N$ . Na prova usamos a seguinte desigualdade clássica devido a Hoeffding [161].

**Lema 2.2.** Seja  $X$  uma variável aleatória com  $a \leq X \leq b$ . Então para qualquer  $s \in \mathbb{R}$ ,

$$\mathbb{E} e^{sX} \leq e^{s\mathbb{E} X + \frac{s^2(b-a)^2}{8}}.$$

A prova está na Seção A.1 do Apêndice.

**Prova do Teorema 2.2.** Primeiro observe que

$$\begin{aligned} \frac{W_n}{W_0} &= \mathbb{E} \sum_{i=1}^N e^{-\gamma L_{i,n}} \leq N \max_{i=1, \dots, N} e^{-\gamma L_{i,n}} \\ &= N \min_{i=1, \dots, N} e^{\gamma L_{i,n}}. \end{aligned} \quad (2.1)$$

Por outro lado, para cada  $t = 1, \dots, n$ ,

$$\begin{aligned} \frac{W_t}{W_{t-1}} &= \mathbb{E} \sum_{i=1}^N e^{-\gamma L_{i,t}} e^{-\gamma L_{i,t-1}} \\ &= \mathbb{E} \sum_{i=1}^N e^{-\gamma L_{i,t-1}} e^{-\gamma L_{i,t}} \\ &= \mathbb{E} \sum_{i=1}^N e^{-\gamma L_{i,t-1}} e^{-\gamma L_{i,t}}. \end{aligned}$$

Agora, usando o Lema 2.2, observamos que a quantidade acima pode ser limitada superiormente por

$$\begin{aligned} &\sum_{i=1}^N e^{-\gamma L_{i,t-1}} e^{-\gamma L_{i,t}} + \frac{\gamma^2}{8} \sum_{i=1}^N e^{-\gamma L_{i,t-1}} e^{-\gamma L_{i,t}} \\ &= \sum_{i=1}^N e^{-\gamma L_{i,t-1}} e^{-\gamma L_{i,t}} + \frac{\gamma^2}{8} \sum_{i=1}^N e^{-\gamma L_{i,t-1}} e^{-\gamma L_{i,t}} \\ &= \sum_{i=1}^N e^{-\gamma L_{i,t-1}} e^{-\gamma L_{i,t}} + \frac{\gamma^2}{8} \sum_{i=1}^N e^{-\gamma L_{i,t-1}} e^{-\gamma L_{i,t}} \end{aligned}$$

onde usamos a convexidade da função de perda em seu primeiro argumento e a definição do predictor de média ponderada exponencialmente. Somando sobre  $t = 1, \dots, n$ , nós obtemos

$$\ln \frac{W_n}{W_0} \leq \frac{\gamma^2}{8} n.$$

Combinando isso com o limite inferior (2.1) e resolvendo para  $L_n$ , descobrimos que

$$L_n \leq \min_{i=1, \dots, N} L_{i,n} + \frac{\gamma^2}{8} n.$$

conforme desejado. ■

### 2.3 Limites que se mantêm uniformemente ao longo do tempo

Como apontamos na seção anterior, o predictor de média ponderada exponencial tem a desvantagem de que o limite de arrependimento do Corolário 2.2 não se mantém uniformemente sobre sequências de qualquer comprimento, mas apenas para sequências de um determinado comprimento  $n$ , onde  $n$  é o valor usado para escolher o parâmetro  $\gamma$ . Para corrigir esse problema, pode-se usar o chamado "truque de duplicação". A ideia é particionar o tempo em períodos de comprimentos exponencialmente crescentes. Em cada período, o predictor de média ponderada é usado com um parâmetro  $\gamma$  escolhido de forma otimizada para o comprimento do intervalo. Quando o período termina, o predictor de média ponderada é reiniciado e então é iniciado novamente no próximo período com um novo valor para  $\gamma$ . Se o truque de duplicação for usado com o predictor de média ponderada exponencialmente, então ele alcança, para qualquer sequência  $y_1, y_2, \dots, y_N$  de resultados e para qualquer  $n \geq 1$ ,

$$L_n \leq \min_{i=1, \dots, N} L_{i,n} + \frac{\gamma^2}{8} n.$$

(ver Exercício 2.8). Este limite é pior que o do Teorema 2.2 por um fator de  $\gamma^2/(\gamma^2 - 1)$ , que é cerca de 3,41.

Considerando que o truque de duplicação redefine os pesos do predictor subjacente após cada período, pode-se perguntar se um limite melhor poderia ser obtido por um argumento mais direto. Na verdade, podemos evitar o truque de duplicação completamente usando o predictor de média ponderada com um potencial que varia no tempo. Ou seja, deixamos o parâmetro  $\gamma$  do potencial exponencial depender do número redondo  $t$ . Como os melhores limites não uniformes para o potencial exponencial são obtidos escolhendo  $\gamma = \gamma^*(\ln N)/n$ , uma escolha natural para um potencial exponencial que varia no tempo é, portanto,  $\gamma_t = \gamma^*(\ln N)/t$ . Ao adaptar a abordagem usada para provar o Teorema 2.2, obtemos para essa escolha de  $\gamma_t$  um limite de arrependimento cujo termo principal é  $2\gamma^*(\ln N)/2 \ln N$  e, portanto, é melhor do que o limite do truque de duplicação. Mais precisamente, provamos o seguinte resultado.

**Teorema 2.3.** *Suponha que a função de perda seja convexa em seu primeiro argumento e assumamos valores em  $[0, 1]$ . Para todo  $n \geq 1$  e para todo  $y_1, \dots, y_n \in Y$ , o arrependimento do predictor de média ponderada exponencialmente com parâmetro de variação temporal  $\gamma_t = \gamma^*(\ln N)/t$  satisfaz*

$$L_n \leq \min_{i=1, \dots, N} L_{i,n} + \frac{\gamma^2}{8} n.$$

$p_t = \prod_{e=1}^{N_{to}} f_i, t, \tilde{y}_1^1 / W_{t, \tilde{y}_1^1}$ , onde  $W_{t, \tilde{y}_1^1} = \prod_{c, j, \tilde{y}_1^1=1}^{N_{to}} e^{w_i, \tilde{y}_1^1} = e^{\sum \tilde{y}_1^1 Li, \tilde{y}_1^1}$  o parâmetro inicial é escolhido como  $\tilde{y}_t = \tilde{y}_a (\ln N)/t$ , onde  $a > 0$  é determinado pela análise. Usamos  $w_i, \tilde{y}_1^1$  para denotar o peso  $w_i, \tilde{y}_1^1$ , onde o parâmetro  $\tilde{y}_t$  é substituído por  $\tilde{y}_t^1$ . Finalmente, usamos  $kt$  para denotar o especialista cuja perda após o primeiro  $t$  rodadas é o menor (empates são desfeitos escolhendo o especialista com menor índice). Ou seja,  $K_{kt, t} = \min_i N_{Li, t}$ . Na prova do teorema, também fazemos uso das seguintes técnicas lema.

$$\text{em} \frac{\overset{\text{Nšo}}{\text{eu}=1} \text{e-} \ddot{\text{y}} \text{di}}{\underset{\text{i}=1}{\text{Nšo}} \text{e} \ddot{\text{y}} \text{y} \text{dj}} \ddot{\text{y}} \frac{\ddot{\text{y}} \ddot{\text{y}} \ddot{\text{y}}}{\text{um}} \text{Em N.}$$
$$em \frac{e_{u=1}^{N_{10}} e^{-\ddot{y}d_i}}{e_{i=1}^{N_{10}} e^{\ddot{y}d_j}} = em \frac{e_{u=1}^{N_{10}} e^{-\ddot{y}d_i}}{e_{i=1}^{N_{10}} e^{(\ddot{y}\ddot{y})d_j} e^{\ddot{y}d_j}} = \ddot{y} \ln E(e^{(\ddot{y}\ddot{y})D}) \ddot{y} (\ddot{y} \ddot{y} \ddot{y}) E D$$
$$\begin{array}{ccc} \text{em } N \ddot{y} \ H(D) & & \\ \text{Nāo} & & \text{Nāo} \\ = & \text{e} \ddot{y} \ddot{y} \text{di } \ddot{y} \text{di} + \ln & \text{e-} \ddot{y} \text{dk} \\ \text{eu}=1 & & \text{k}=1 \\ & \text{Nāo} & \\ = \ddot{y} \text{E } D + \ln & & \text{e-} \ddot{y} \text{dk} \\ & \text{k}=1 & \\ \ddot{y} \ \ddot{y} \text{E } D, & & \end{array} \quad \begin{array}{c} 1 \\ \hline \text{Nāo} \\ \text{j}=1 \end{array} \quad \begin{array}{c} \ddot{y} \ddot{y} \text{dj} \end{array}$$

**Prova do Teorema 2.3.** Assim como na prova do Teorema 2.2, estudamos a evolução de  $\ln(Wt/Wt\tilde{y}1)$ . No entanto, aqui precisamos acoplar isso com  $\ln(wkt\tilde{y}1,\tilde{t}y1/wkt,t)$ , incluindo em ambos os termos o parâmetro de variação temporal  $\tilde{y}t$ . Acompanhando o melhor especialista atual,  $kt$  é usado para limitar o peso  $\ln(wkt,t/Wt)$ . Na verdade, o peso do melhor especialista geral (após  $n$  rodadas) poderia ficar arbitrariamente pequeno durante o processo de previsão. Assim, escrevemos o

seguinte:

$$\frac{1}{ot} \text{ em } \frac{wkt\bar{y}_1, \bar{t}\bar{y}_1}{\text{Peso-1}} = \frac{1}{\bar{y}_{t+1}} \text{ em } \frac{\text{trabalho}, t}{\text{Peso}} + \frac{1}{\bar{y}_{t+1}} \text{ em } \frac{\text{peso corporal}, t/\text{peso corporal}}{\text{peso}, t/\text{peso}} + \frac{1}{\bar{y}_{t+1}} \text{ em } \frac{wkt\bar{y}_1, \bar{t}\bar{y}_1/Wt\bar{y}_1}{\text{peso corporal}, t/\text{peso corporal}}.$$

(UM) (B) (C)

Agora, ligamos separadamente os três termos do lado direito. O termo (A) é facilmente limitado pela observação de que  $\bar{y}_{t+1} < \bar{y}_t$  e usando o fato de que  $kt$  é o índice do especialista com a menor perda após as primeiras  $t$  rodadas. Portanto,  $wkt, t/Wt$  deve ser pelo menos  $1/N$ . Assim, nós ter

$$(Um) = \frac{1}{\bar{y}_{t+1}} \frac{1}{ot} \text{ em } \frac{\text{Peso}}{wkt, t} \bar{y} = \frac{1}{\bar{y}_{t+1}} \frac{1}{ot} \text{ Em } N.$$

Procedemos à delimitação do termo (B) da seguinte forma:

$$(B) = \frac{1}{ot} \text{ em } \frac{\text{peso corporal}, t/\text{peso corporal}}{\text{peso}, t/\text{peso}} = \frac{1}{ot} \text{ em } \frac{\sum_{j=1}^{N_{\text{Não}}} e^{\bar{y}_{j,t+1}(L_{i,t} \bar{y}_{Lkt}, t)}}{\sum_{j=1}^{N_{\text{Não}}} e^{\bar{y}_{j,t}(L_{j,t} \bar{y}_{Lkt}, t)}} \\ \bar{y} \frac{\bar{y}_t \bar{y}_{t+1}}{\bar{y}_{t+1}} \text{ em } N = \frac{1}{\bar{y}_{t+1}} \frac{1}{ot} \text{ em } N,$$

onde a desigualdade é provada pela aplicação do Lema 2.3 com  $d_i = L_{i,t} \bar{y}_{Lkt+1, t}$ . Note que  $d_i \geq 0$  porque  $kt$  é o índice do especialista com a menor perda após as primeiras  $t$  rodadas e  $\sum_{i=1}^{N_{\text{Não}}} e^{\bar{y}_{i,t+1} d_i} \geq 1$  como para  $i = kt+1$  temos  $d_i = 0$ . O termo (C) é primeiro dividido como segue:

$$(C) = \frac{1}{ot} \text{ em } \frac{wkt\bar{y}_1, \bar{t}\bar{y}_1/Wt\bar{y}_1}{\text{peso corporal}, t/\text{peso corporal}} = \frac{1}{ot} \text{ em } \frac{wkt\bar{y}_1, \bar{t}\bar{y}_1}{\text{peso}, t} + \frac{1}{ot} \text{ em } \frac{C}{\text{Peso-1}}.$$

Tratamos separadamente cada um dos dois subtermos do lado direito. Para o primeiro, nós temos

$$\frac{1}{ot} \text{ em } \frac{wkt\bar{y}_1, \bar{t}\bar{y}_1}{\text{peso}, t} = \frac{1}{ot} \text{ em } \frac{e^{\bar{y}_{kt, t} L_{kt, t} \bar{y}_1}}{e^{\bar{y}_{kt, t} L_{kt, t}}} = L_{kt, t} \bar{y} L_{kt, t} \bar{y}_1.$$

Para o segundo subtermo, procedemos de forma semelhante à prova do Teorema 2.2, aplicando Limite de Hoeffding (Lema 2.2) para a variável aleatória  $Z$  que assume o valor  $(f_i, t, y_t)$  com probabilidade  $w_i, \bar{t}\bar{y}_1/Wt\bar{y}_1$  para cada  $i = 1, \dots, N$ :

$$\frac{1}{ot} \text{ em } \frac{C}{\text{Peso-1}} = \frac{1}{\text{em } \bar{y}_t} \sum_{i=1}^{N_{\text{Não}}} \frac{w_i, \bar{t}\bar{y}_1}{\text{Peso-1}} e^{\bar{y}_{i,t} (f_i, t, y_t)} \\ \bar{y} \bar{y} \sum_{i=1}^{N_{\text{Não}}} \frac{w_i, \bar{t}\bar{y}_1}{\text{Peso-1}} (f_i, t, y_t) + \frac{ot}{8} \\ \bar{y} \bar{y} (\text{ponto}, \text{ano}) + \frac{ot}{8},$$

onde na última etapa usamos a convexidade da perda. Finalmente, substituímos de volta na equação principal os limites dos dois primeiros termos (A) e (B), e os limites dos dois

$$\begin{aligned} & (pt, yt) \tilde{y} Lkt, t \tilde{y} Lkt\tilde{y}1, t\tilde{y}1 + \tilde{y}t \\ & \frac{\tilde{y}t \ln N}{8} \frac{1}{1} \\ & + \frac{1}{Peso \tilde{y}t+1} \frac{em}{trabalho, t} \frac{1 \text{ semana} \tilde{y}1, t\tilde{y}1 \ln}{\tilde{y}t \quad Peso \tilde{y}1} \\ & + 2 \frac{1}{\tilde{y}t+1} \frac{1}{ot} \quad Em N. \end{aligned}$$
$$\begin{array}{l} \text{rdo} \\ \text{t=1} \end{array} \quad \frac{1}{\text{yt}+1} \text{ em } \frac{\text{trabalho.t}}{\text{Peso}} \quad \frac{1}{\text{ot}} \text{ em } \frac{\text{wkt}\bar{y}_1.\bar{t}\bar{y}_1}{\text{Peso}-1} \quad \bar{y} \ \bar{y} \ \bar{y}_1 \quad \frac{1}{\text{em}} \text{ em } \frac{\text{semana0.0}}{\text{W0}} = \frac{\text{em } N}{\text{um}}$$

$$\begin{array}{l} \text{rdo} \\ \text{t=1} \end{array} \quad \frac{1}{\text{yt}+1} \quad \frac{1}{\text{ot}} = \frac{n+1}{a(\text{em } N)} \quad \frac{1}{\text{um}(\text{em } N)}$$
$$L_n \text{ } \ddot{y} \text{ min i=1,...,N} \quad L_{i,n+2} \frac{\ddot{y}_{um \text{ em } N}}{4} \quad \frac{(n+1) \text{ em } N}{um} \quad \frac{\text{em } N}{um}$$
$$Ln \tilde{y} \min_{i=1, \dots, N} \quad Li, n+2 \quad \frac{n+1}{2} \text{ em } N+ \quad \frac{n+1}{8} \text{ em } N$$

■

Para entender por que se pode esperar melhores limites para o arrependimento quando a perda cumulativa do melhor especialista é pequena, lembre-se do exemplo simples descrito na introdução quando  $Y = D = \{0, 1\}$  e  $(p, y) = |p \tilde{y} - y| \tilde{y} \in \{0, 1\}$  (este exemplo viola nossa suposição de que  $D$  é um conjunto convexo, mas ajuda a entender o fenômeno básico). Se o previsor souber com antecedência que um dos  $N$  especialistas sofrerá perda zero, ou seja,  $\min_i L_i = 0$ , então ele pode prever usando o seguinte simples “voto majoritário”. No tempo  $t = 1$ , preveja  $p_1 = 0$  se e somente se pelo menos metade dos  $N$  especialistas preverem 0. Após o primeiro bit  $y_1$  ser revelado, descarte todos os especialistas com  $f_{i,1} = y_1$ . No tempo  $t = 2$ , preveja  $p_2 = 0$  se e somente se pelo menos metade dos  $N$  especialistas restantes preverem 0, descarte todos os especialistas que previram incorretamente após  $y_2$  ser revelado, e assim por diante. Portanto, cada vez que o previsor comete um erro, pelo menos metade dos especialistas sobreviventes são descartados (porque o previsor sempre vota de acordo com a maioria dos especialistas restantes). Se apenas um especialista permanecer, o previsor não faz mais nenhuma

erro. Assim, o número total de erros do previsor (que, neste caso, coincide com seu arrependimento) é no máximo

$\log_2 N$  Nesta seção mostramos que limites de arrependimento da forma  $O(\ln N)$  são possíveis para todos perdas limitadas e convexas em espaços de decisão convexos sempre que o previsor recebe a informação de que o melhor especialista não sofrerá nenhuma perda. Note que tais limites são significativamente melhor que  $\tilde{y}(n/2)$  em  $N$ , que é válido independentemente da perda do melhor especialista.

Para simplificar, escrevemos  $L_{\tilde{y}} = \min_{i=1, \dots, N} L_i$ . Agora mostramos que sempre que  $L_{\tilde{y}}$  é conhecido de antemão, o parâmetro  $\tilde{y}$  do previsor de média ponderada exponencialmente pode ser escolhido de modo que seu arrependimento é limitado por  $2L_{\tilde{y}} \ln N + \ln N$ , que é igual a  $\ln N$  quando  $L_{\tilde{y}} = 0$  e é da ordem  $\tilde{y} n \ln N$  quando  $L_{\tilde{y}}$  é de ordem  $n$ . Nossa principal ferramenta é o seguinte refinamento de Teorema 2.2.

**Teorema 2.4.** *Suponha que a função de perda seja convexa em seu primeiro argumento e que ela assume valores em  $[0, 1]$ . Então, para qualquer  $\tilde{y} > 0$ , o arrependimento da média ponderada exponencialmente o meteorologista satisfaz*

$$L_n \tilde{y} \leq \frac{\tilde{y} L_{\tilde{y}} + \ln N}{1 - e^{-\tilde{y}}}.$$

É fácil perceber que, em alguns casos, uma escolha desinformada de  $\tilde{y}$  ainda pode levar a um bom arrependimento vinculado.

**Corolário 2.3.** *Suponha que o previsor de média ponderada exponencialmente seja usado com  $\tilde{y} = 1$ . Então, sob as condições do Teorema 2.4,*

$$L_n \tilde{y} \leq \frac{e}{e - 1} L_{\tilde{y}} + \ln N.$$

Este limite é muito melhor do que o limite geral do Teorema 2.2 se  $L_{\tilde{y}} = \tilde{y} n$ , mas pode ser muito pior caso contrário.

Agora derivamos um novo limite ajustando  $\tilde{y}$  no Teorema 2.4 em termos da perda total  $L_{\tilde{y}}$  de o melhor especialista.

**Corolário 2.4.** *Suponha que o previsor de média ponderada exponencial seja usado com  $\tilde{y} = \frac{\ln N}{L_{\tilde{y}}}$  em  $1 + (2 \ln N)/L_{\tilde{y}}$ , onde  $L_{\tilde{y}} > 0$  deve ser conhecido com antecedência. Então, sob o condições do Teorema 2.4,*

$$L_n \tilde{y} \leq \frac{2L_{\tilde{y}}}{\ln N} \ln N + \ln N.$$

**Prova.** Usando o Teorema 2.4, precisamos apenas mostrar que, para nossa escolha de  $\tilde{y}$ ,

$$\frac{\ln N + \tilde{y} L_{\tilde{y}}}{N \cdot 1 - e^{-\tilde{y}}} \leq e^{-\tilde{y}} L_{\tilde{y}} + \ln N + 2L_{\tilde{y}} \ln N \quad (2.2)$$

Partimos da desigualdade elementar  $(e^{\tilde{y}} - e^{-\tilde{y}})/2 = \sinh(\tilde{y}) \geq \tilde{y}$ , que vale para todos  $\tilde{y} \geq 0$ . Substituindo o  $\tilde{y}$  no numerador do lado esquerdo de (2.2) por este limite superior, descobrimos que (2.2) está implícito em

$$\frac{\ln N}{1 - e^{-\tilde{y}}} + \frac{1 + e^{-\tilde{y}}}{2e^{\tilde{y}} - 1} L_{\tilde{y}} \leq e^{-\tilde{y}} L_{\tilde{y}} + \ln N + 2L_{\tilde{y}} \ln N.$$

A prova é concluída observando que a desigualdade acima é válida para a igualdade para o nosso escolha de  $\tilde{y}$ . ■

É claro que a quantidade  $L_{\tilde{y}}^n$  só está disponível após a  $n$ -ésima rodada de previsão. A falta desta informação pode ser compensada permitindo que  $\tilde{y}$  mude de acordo com a perda da atualmente o melhor especialista, de forma semelhante à mostrada na Seção 2.3 (ver Exercício 2.10). o limite de arrependimento que pode ser obtido por meio dessa abordagem é da forma  $2L\tilde{y} \ln N + c \ln N$ , onde  $c > 1$  é uma constante. Note que, similarmente ao Teorema 2.3, o uso de um parâmetro que varia com o tempo  $\tilde{y}_t$  leva a um limite cujo termo principal é o dobro daquele obtido quando  $\tilde{y}$  é fixo e escolhido otimamente com base no horizonte  $n$  (como no Teorema 2.2) ou na perda  $L_{\tilde{y}}^n$  dos melhores especialista (como no Corolário 2.4).

**Prova do Teorema 2.4.** A prova é uma modificação simples da do Teorema 2.2. A única diferença é que a desigualdade de Hoeffding agora é substituída pelo Lema A.3 (veja o Apêndice). Lembre-se da prova do Teorema 2.2 que

$$L_{\tilde{y}}^n = \sum_{t=1}^n \ell_t(\tilde{y}_t) = \sum_{t=1}^n \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t) = \sum_{t=1}^n \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t) \mathbb{E} \ell_t(i, \tilde{y}_t) + \sum_{t=1}^n \sum_{i=1}^N w_{i,t} (\ell_t(i, \tilde{y}_t) - \mathbb{E} \ell_t(i, \tilde{y}_t)).$$

Aplicamos o Lema A.3 à variável aleatória  $X_t$  que assume valor  $\ell_t(i, \tilde{y}_t)$  com probabilidade  $w_{i,t}/W_t$  para cada  $i = 1, \dots, N$ . Note que pela convexidade da função de perda e pela equação de Jensen desigualdade,  $\mathbb{E} \ell_t(\tilde{y}_t) \leq \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t)$  e portanto, pelo Lema A.3,

$$\sum_{t=1}^n \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t) \leq \sum_{t=1}^n \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t) \mathbb{E} \ell_t(i, \tilde{y}_t) + \sum_{t=1}^n \sum_{i=1}^N w_{i,t} (\ell_t(i, \tilde{y}_t) - \mathbb{E} \ell_t(i, \tilde{y}_t)) \leq \sum_{t=1}^n \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t) \mathbb{E} \ell_t(i, \tilde{y}_t) + \sum_{t=1}^n \sum_{i=1}^N w_{i,t} (\ell_t(i, \tilde{y}_t) - \mathbb{E} \ell_t(i, \tilde{y}_t)).$$

Por isso,

$$L_{\tilde{y}}^n = \sum_{t=1}^n \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t) \leq \sum_{t=1}^n \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t) \mathbb{E} \ell_t(i, \tilde{y}_t) + \sum_{t=1}^n \sum_{i=1}^N w_{i,t} (\ell_t(i, \tilde{y}_t) - \mathbb{E} \ell_t(i, \tilde{y}_t)) \leq \sum_{t=1}^n \sum_{i=1}^N w_{i,t} \ell_t(i, \tilde{y}_t) \mathbb{E} \ell_t(i, \tilde{y}_t) + \sum_{t=1}^n \sum_{i=1}^N w_{i,t} (\ell_t(i, \tilde{y}_t) - \mathbb{E} \ell_t(i, \tilde{y}_t)).$$

Resolver para  $L_n$  produz o resultado. ■

## 2.5 Previsores usando o gradiente da perda

Considere novamente o predictor de média ponderada exponencial cujas previsões são definidas por

$$p_t = \frac{\sum_{i=1}^N w_{i,t} \ell_t(i, p_t)}{\sum_{i=1}^N w_{i,t}},$$

onde o peso  $w_{i,t}$  para o especialista  $i$  na rodada  $t$  é definido por  $w_{i,t} = e^{-\tilde{y}_t L_{i,t}}$ . Nesta seção introduzimos e analisamos um predictor de média ponderada exponencialmente diferente no qual a perda cumulativa  $L_{i,t}$  que aparece no expoente de  $w_{i,t}$  é substituída pelo gradiente da perda somada ao tempo  $t-1$ . Esta nova classe de previsores será generalizada em Capítulo 11, onde também fornecemos ampla análise e motivação.

Lembre-se de que o espaço de decisão  $D$  é um subconjunto convexo de um espaço linear. Ao longo deste seção, também assumimos que  $D$  é de dimensão finita, embora essa suposição possa ser relaxada facilmente. Se for diferenciável, usamos  $\nabla \ell(p, y)$  para denotar seu gradiente em relação ao primeiro argumento  $p \in D$ .



Defina o *previsor de média ponderada exponencialmente baseado em gradiente* por

$$p_t = \frac{\sum_{i=1}^N \exp(-\sum_{s=1}^{t-1} \tilde{y}_i(p_s, y_s) \cdot f_{i,s})}{\sum_{j=1}^N \exp(-\sum_{s=1}^{t-1} \tilde{y}_j(p_s, y_s) \cdot f_{j,s})}.$$

Para entender a intuição por trás desse preditor, observe que o peso atribuído ao especialista  $i$  é pequeno se a soma dos produtos internos  $\tilde{y}_i(p_s, y_s) \cdot f_{i,s}$  foi grande no passado. Esse produto interno é grande se o conselho do especialista  $i$  aponta na direção do maior aumento da função de perda. Um valor tão grande significa que ter atribuído um peso um pouco maior a esse especialista teria aumentado a perda sofrida no tempo  $s$ . De acordo com a filosofia do previsor de média ponderada exponencialmente baseado em gradiente, o peso de tal especialista tem que ser diminuído.

As previsões deste previsor são, é claro, geralmente diferentes daquelas do previsor de média ponderada exponencial padrão. No entanto, observe que no caso especial de previsão binária com perda absoluta (ou seja, se  $D = [0, 1]$ ,  $Y = \{0, 1\}$  e  $\langle x, y \rangle = |x - y|$ ), uma configuração que estudamos em detalhes no Capítulo 8, as previsões dos dois previsores são idênticas (veja os exercícios).

Agora mostramos que, sob condições adequadas na norma do gradiente da perda, o arrependimento do novo previsor pode ser limitado pela mesma quantidade que foi usada para limitar o arrependimento do previsor padrão de média ponderada exponencialmente no Corolário 2.2.

**Corolário 2.5.** *Suponha que o espaço de decisão  $D$  seja um subconjunto convexo da bola unitária euclidiana  $\{q \in \mathbb{R}^d : \|q\| \leq 1\}$ , a função de perda seja convexa em seu primeiro argumento e que seu gradiente  $\tilde{y}$  exista e satisfaça  $\|\tilde{y}\| \leq 1$ . Para qualquer  $n$  e  $\gamma > 0$ , e para todos  $y_1, \dots, y_n \in Y$ , o arrependimento do previsor de média ponderada exponencial baseado em gradiente satisfaz*

$$L_n \leq \min_{e \in \{1, \dots, N\}} L_{i,n} + 2 \frac{\exp(N\gamma)}{\gamma}.$$

**Prova.** O vetor de peso  $w_t = (w_{t,1}, \dots, w_{t,N})$  usado por este previsor tem componentes

$$w_{i,t} = \exp\left(-\sum_{s=1}^{t-1} \tilde{y}_i(p_s, y_s) \cdot f_{i,s}\right).$$

Observe que esses pesos correspondem à média ponderada exponencial do previsor com base na função de perda definida no tempo  $t$  por

$$(q, y_t) = q \cdot \tilde{y}(p_t, y_t), \quad q \in D.$$

Por suposição, assume valores em  $[0, 1]$ . Aplicando o Teorema 2.2 após redimensionar em  $[0, 1]$  (ver Seção 2.6), obtemos

$$\max_{i=1, \dots, N} \sum_{t=1}^n (p_t \cdot \tilde{y}_i(f_t, t) - \tilde{y}(p_t, y_t)) = \min_{e \in \{1, \dots, N\}} \sum_{t=1}^n (p_t \cdot \tilde{y}_e(f_t, t) - \tilde{y}(p_t, y_t)) + 2 \frac{\exp(N\gamma)}{\gamma}.$$

A prova é completada expandindo  $(f_i, t, y_t)$  em torno de  $(p_t, y_t)$  como segue:

$$(p_t, y_t) \tilde{y} (f_i, t, y_t) \tilde{y} (p_t \tilde{y} f_i, t) \cdot \tilde{y} (p_t, y_t),$$

o que implica que

$$L_n \tilde{y} \min_{e_u=1, \dots, N} L_{i,n} \tilde{y} \min_{t=1}^n (p_t, y_t) \tilde{y} \min_{e_u=1, \dots, N} (f_i, t, y_t). \quad \blacksquare$$

## 2.6 Perdas Escaladas e Jogos Assinados

Até este ponto, sempre assumimos que o intervalo da função de perda é o intervalo unitário  $[0, 1]$ . Agora investigamos como as escalas e traduções deste intervalo afetam as estratégias de previsão e seu desempenho.

Considere primeiro o caso de uma função de perda  $\tilde{y} \in [0, M]$ . Se  $M$  for conhecido, podemos executar o predictor de média ponderada nas perdas escalonadas  $\tilde{y}/M$  e aplicar sem nenhuma modificação a análise desenvolvida para o caso  $[0, 1]$ . Por exemplo, no caso do predictor de média ponderada exponencialmente, o Corolário 2.4, aplicado a essas perdas escalonadas, produz o limite de arrependimento

$$L_n \tilde{y} L_{\tilde{y}} \tilde{y} 2L_{\tilde{y}} M \ln N + M \ln N.$$

O termo aditivo  $M \ln N$  é necessário. De fato, se é tal que para todo  $p \in D$  existem  $p \in D$  e  $y \in Y$  tais que  $(p, y) = M$  e  $(p, y) = 0$ , então o conselho do especialista pode ser escolhido de modo que qualquer predictor incorra em uma perda cumulativa de pelo menos  $M \log N$  em alguma sequência de resultados com  $L_{\tilde{y}} = 0$ .

Considere agora o intervalo traduzido  $[\tilde{y}M, 0]$ . Se interpretarmos perdas negativas como ganhos, podemos introduzir o arrependimento  $G_{\tilde{y}} \tilde{y} G_n$  medindo a diferença entre o ganho cumulativo  $G_{\tilde{y}} = \tilde{y}L_{\tilde{y}}$   $G_n = \max_{i=1, \dots, N} (\tilde{y}L_{i,n})$  do melhor especialista e o ganho cumulativo  $G$  do predictor.  $G_{\tilde{y}} = \tilde{y}L_n$  de Como antes, se  $M$  for conhecido, podemos executar o predictor de média ponderada nos ganhos escalonados  $(\tilde{y})/M$  e aplicar a análise desenvolvida para funções de perda com valores  $[0, 1]$ . Adaptando o Corolário 2.4 obtemos um limite da forma

$$G_n \tilde{y} G_{\tilde{y}} \tilde{y} 2G_{\tilde{y}} M \ln N + M \ln N.$$

Observe que o arrependimento agora é dimensionado com o maior ganho cumulativo  $G_{\tilde{y}}$ .

Agora nos voltamos para o caso geral em que os predictors são pontuados usando uma *função de recompensa genérica*  $h : D \times Y \rightarrow \mathbb{R}$ , côncava em seu primeiro argumento. O objetivo do predictor é maximizar sua recompensa cumulativa. O arrependimento correspondente é definido por

$$\max_{e_u=1, \dots, N} \sum_{t=1}^n h(f_i, t, y_t) \tilde{y} \sum_{t=1}^n h(p_t, y_t) = H_{\tilde{y}} - H_n.$$

Se a função de payoff  $h$  tem alcance  $[0, M]$ , então é uma função de ganho e o predictor joga um *jogo de ganho*. Similarmente, se  $h$  tem alcance em  $[\tilde{y}M, 0]$ , então é o negativo de uma função de perda e o predictor joga um *jogo de perda*. Finalmente, se o alcance de  $h$  inclui uma vizinhança de 0, então o jogo jogado pelo predictor é um *jogo assinado*.

Traduzido para esta terminologia, os argumentos propostos no início desta seção dizem que em qualquer jogo não assinado (ou seja, qualquer jogo de perda  $[\tilde{y}M, 0]$  ou jogo de ganho  $[0, M]$ ),

o reescalamento dos payoffs produz um limite de arrependimento de ordem  $\frac{H\tilde{y}}{M} \ln N$  sempre que  $M$  é conhecido. No caso de jogos assinados, no entanto, o escalonamento não é suficiente. De fato, se  $h \in [\tilde{y}M, M]$ , então a redução para o caso  $[0, 1]$  é obtida pela transformação linear  $h \mapsto (h + M)/(2M)$ . Aplicando isso à análise que leva ao Corolário 2.4, obtemos o limite de arrependimento

$$H^* - H_{\text{mult}} \leq \frac{H\tilde{y}}{M} \ln N + 2M \text{ em } N = O\left(\frac{H\tilde{y}}{M} \ln N\right)$$

Isso mostra que, para jogos assinados, reduzir para o caso  $[0, 1]$  pode não ser a melhor coisa a fazer.

Idealmente, gostaríamos de substituir o fator  $n$  no termo principal por algo como  $|h(f_i, 1, y_1)| + \dots + |h(f_i, n, y_n)|$  para um especialista arbitrário  $i$ . Nas próximas seções mostramos que, em certos casos, podemos fazer ainda melhor do que isso.

## 2.7 O Previsor Multilinear

Funções potenciais oferecem uma ferramenta conveniente para derivar previsores de média ponderada. No entanto, bons previsores para jogos assinados também podem ser projetados sem usar potenciais, como mostrado nesta seção.

Corrija um jogo assinado com função de pagamento  $h: D \times Y \rightarrow \mathbb{R}$  e considere a ponderação previsor médio que prevê, no tempo  $t$ ,

$$p_t = \frac{\sum_{i=1}^N w_{i,t} \tilde{y}_i f_{i,t}}{\sum_{i=1}^N w_{i,t}},$$

onde  $w_{i,t}$  são os pesos

Os pesos  $w_{i,t}$  deste previsor são definidos recursivamente  $w_{j,t} = 1$  para  $j=1$

como segue

$$w_{i,t} = \begin{cases} 1 & \text{se } t = 0 \\ w_{i,t-1} + \tilde{y}_i h(f_i, t, y_t) & \text{caso contrário,} \end{cases}$$

onde  $\tilde{y} > 0$  é um parâmetro do previsor. Como  $w_{i,t}$  é uma forma multilinear dos payoffs, chamamos isso de *previsor multilinear*.

Note que os pesos  $w_{1,t}, \dots, w_{N,t}$  não podem ser expressos como funções do arrependimento. Por outro lado, como  $(1 + \tilde{y}h) \leq e^{\tilde{y}h}$ , o arrependimento do previsor multilinear pode ser limitado por meio de uma técnica similar à usada na prova do Teorema 2.2 para o previsor de média ponderada exponencialmente. Precisamos apenas do seguinte lema simples (a prova é deixada como exercício).

**Lema 2.4.** Para todo  $z \geq 1/2$ ,  $\ln(1 + z) \leq z \leq z^2$ .

O próximo resultado mostra que o arrependimento do previsor multilinear é naturalmente expresso em termos da soma quadrada dos ganhos de um especialista arbitrário.

**Teorema 2.5.** Suponha que a função de payoff  $h$  seja côncava em seu primeiro argumento e satisfaça  $h \in [\tilde{y}M, M]$ . Para qualquer  $n$  e  $0 < \tilde{y} < 1/(2M)$ , e para todos  $y_1, \dots, y_n \in Y$ , o arrependimento do previsor multilinear satisfaz

$$O\left(\frac{H\tilde{y}}{M} \ln N + \sum_{t=1}^n \frac{1}{\tilde{y}} \sum_{i=1}^N w_{i,t} h(f_i, t, y_t)\right) \leq \frac{1}{\tilde{y}} \sum_{t=1}^n \sum_{i=1}^N w_{i,t} h(f_i, t, y_t)^2 \quad \text{para cada } i = 1, \dots, N.$$

**Prova.** Para qualquer  $i = 1, \dots, N$ , note que  $h(f_i, t, y_t) \leq \tilde{y}M$  e  $\tilde{y} \leq 1/(2M)$  implicam que  $\tilde{y}h(f_i, t, y_t) \leq \tilde{y}^2/2$ . Portanto, podemos aplicar o Lema 2.4 a  $\tilde{y}h(f_i, t, y_t)$  e obter

$$\begin{aligned} \frac{W_n}{W_0} &= \tilde{y} \sum_{t=1}^n \left( 1 + \tilde{y}h(f_i, t, y_t) \right) \\ &= \tilde{y} \sum_{t=1}^n 1 + \sum_{t=1}^n \tilde{y}h(f_i, t, y_t) \\ &\leq \tilde{y} \sum_{t=1}^n 1 + \sum_{t=1}^n \tilde{y}^2/2 \\ &= \tilde{y} \sum_{t=1}^n 1 + \sum_{t=1}^n \tilde{y}^2 h(f_i, t, y_t) \end{aligned}$$

Por outro lado,

$$\begin{aligned} \frac{W_n}{W_0} &= \sum_{t=1}^n \frac{\text{Peso}_t}{\text{Peso}_0} \\ &= \sum_{t=1}^n \exp \left( \sum_{s=1}^t \tilde{y}h(f_i, s, y_s) \right) \\ &\leq \sum_{t=1}^n \exp \left( \sum_{s=1}^t \tilde{y}h(f_i, s, y_s) \right) \quad (\text{já que } \ln(1+x) \leq x \text{ para } x > -1) \\ &\leq \sum_{t=1}^n \exp \left( \sum_{s=1}^t \tilde{y}H \right) \quad (\text{já que } h(\cdot, y) \text{ é côncavo}). \end{aligned}$$

Combinando os limites superior e inferior de  $\ln(W_n/W_0)$  e dividindo por  $\tilde{y} > 0$ , obtemos o seguinte lema.

Seja  $Q_{\tilde{y}} = h(f_{k,1}, y_1) + \dots + h(f_{k,n}, y_n)$  onde  $k$  é tal que  $H_{k,n} = H_{\tilde{y}}^*$ . Se  $\tilde{y}$  for escolhido usando  $Q_{\tilde{y}}$ , então o Teorema 2.5 implica diretamente o seguinte.

**Corolário 2.6.** Suponha que o predictor multilinear seja executado com

$$\tilde{y} = \min \left\{ \frac{1}{2M}, \frac{1}{Q_{\tilde{y}}} \right\},$$

onde  $Q_{\tilde{y}} > 0$  é suposto ser conhecido antecipadamente. Então, sob as condições do Teorema 2.5,

$$H^* \leq \tilde{y} H + 2 Q_{\tilde{y}} \leq \tilde{y} H + 4M \tilde{y} \leq N.$$

Para apreciar este resultado, considere um jogo de perdas com  $h \in [\tilde{y}M, 0]$  e deixe  $L_{\tilde{y}} = \tilde{y} \max_i H_{i,n}$ .

Como  $Q_{\tilde{y}} \leq M L_{\tilde{y}}$ , a garantia de desempenho do predictor multilinear é no máximo um fator

de  $\tilde{y}^2$  maior do que o do predictor com média ponderada exponencial, cujo arrependimento em

este caso tem o termo principal  $2L_{\tilde{y}} \ln N$  (ver Seção 2.4). Entretanto, em alguns casos  $Q_{\tilde{y}}$

pode ser significativamente menor que  $M L_{\tilde{y}}$ , de modo que o limite do Corolário 2.6 apresenta uma real

melhoria. Na Seção 2.8, mostramos que uma análise mais cuidadosa do exponencialmente

o prognosticador de média ponderada produz um arrependimento de segunda ordem semelhante (embora não comparável) limites.

Ainda é um problema em aberto obter limites de arrependimento de ordem  $\tilde{y}$  Em  $N$  sem explorar algum conhecimento prévio sobre a sequência  $y_1, \dots, y_n$  (ver Exercício 2.14). Na verdade, a análise de técnicas de ajuste adaptativo, como o truque de duplicação ou o  $\tilde{y}$  variável no tempo, dependem a monotonicidade da quantidade cuja evolução determina a estratégia de afinação. por outro lado, a sequência  $\tilde{y}_1, \tilde{y}_2, \dots$  não é necessariamente monótono como  $\tilde{y}$  e  $\tilde{y}_{t+1}$  não pode geralmente estão relacionados quando os especialistas alcançam os maiores retornos cumulativos nas rodadas  $t$  e  $t+1$  são diferentes.

## 2.8 O Previsor Exponencial para Jogos Assinados

Uma ligeira modificação da nossa análise anterior é suficiente para mostrar que o exponencialmente o previsor de média ponderada também é capaz de obter um pequeno arrependimento em jogos assinados. Como o previsor multilinear da Seção 2.7, este novo limite é expresso em termos de somas de termos quadráticos que estão relacionados com a variância das perdas dos especialistas em relação ao distribuição induzida pelos pesos do previsor. Além disso, o uso de uma distribuição variável no tempo potencial permite-nos dispensar a necessidade de qualquer conhecimento prévio das melhores retorno acumulado  $H_{\tilde{y}} = \sum_{t=1}^n H_{i,t} \tilde{y}_t$ .

Começamos redefinindo, para a configuração onde as funções de payoff são usadas, o exponencialmente previsor de média ponderada com potencial de variação temporal introduzido na Seção 2.3. Dado um função de recompensa  $h$ , este previsor prevê com  $p_t = f_{i,t} w_{i,t} \tilde{y}_1 / W_t \tilde{y}_1$  onde  $W_t \tilde{y}_1 = \sum_{i=1}^I w_{i,t} \tilde{y}_1$ ,  $w_{i,t} \tilde{y}_1 = e^{\tilde{y}_t H_{i,t} \tilde{y}_1}$ ,  $H_{i,t} \tilde{y}_1 = h(f_{i,t}, y_1) + \dots + h(f_{i,t}, y_t \tilde{y}_1)$ , e assumimos que a sequência  $\tilde{y}_1, \tilde{y}_2, \dots$  de parâmetros é positiva. Note que o valor de  $\tilde{y}_1$  é irrelevante porque  $H_{i,0} = 0$  para todo  $i$ .

Sejam  $X_1, X_2, \dots$  variáveis aleatórias tais que  $X_t = h(f_{i,t}, y_t)$  com probabilidade  $w_{i,t} \tilde{y}_1 / W_t \tilde{y}_1$  para todos  $i$  e  $t$ . O próximo resultado, cuja prova é deixada como um exercício, limita o arrependimento do previsor exponencial para qualquer sequência não crescente de parâmetros potenciais em termos do processo  $X_1, \dots, X_n$ . Note que este lema não assume nenhuma condição de limitação na função de payoff.

**Lema 2.5.** *Seja  $h$  uma função de payoff côncava em seu primeiro argumento. A função exponencialmente previsor de média ponderada, executado com qualquer sequência não crescente  $\tilde{y}_1, \tilde{y}_2, \dots$  de parâmetros satisfaz, para qualquer  $n \geq 1$  e para qualquer sequência  $y_1, \dots, y_n$  de resultados,*

$$H_{\tilde{y}}^* - H_{\tilde{y}} \leq \sum_{t=1}^n \frac{2}{\tilde{y}_{t+1}} + \frac{1}{\tilde{y}_1} \text{ em } N + \sum_{t=1}^n \frac{1}{\tilde{y}_t} \text{ Em } E e^{\tilde{y}_t (X_t - E X_t)}.$$

Deixar

$$V_t = \sum_{s=1}^n \text{var}(X_s) = \sum_{s=1}^n E X_s^2 - (E X_s)^2.$$

Nosso próximo resultado mostra que, com uma escolha apropriada da sequência  $\tilde{y}_t$ , o arrependimento de o previsor exponencial no tempo  $n$  é no máximo da ordem  $\tilde{y} \sqrt{n} \ln N$ . Note, entretanto, que o bound não está em forma fechada, pois  $\tilde{y}_n$  depende dos pesos do previsor  $w_{i,t}$  para todos os  $i$  e  $t$ .

**Teorema 2.6.** Seja  $h$  uma função de payoff com valor  $[\tilde{y}M, M]$  côncava em seu primeiro argumento.

Suponha que o previsor de média ponderada exponencialmente seja executado com

$$\tilde{y}_t = \min_{\tilde{y}} \frac{1}{2M} \left( \frac{2\tilde{y}_2 \tilde{y}_1}{e^{\tilde{y}_2}} \right)^{\frac{1}{\tilde{y}_1}} \quad \text{em } N, \quad t = 2, 3, \dots$$

Então, para qualquer  $n \geq 1$  e para qualquer sequência  $y_1, \dots, y_n$  de resultados,

$$H_n - H_{\tilde{y}} \leq 4 \sqrt{n \ln N} + 4M \ln N + (e^{\tilde{y}_2} - 2)M.$$

**Prova.** Para resumir, escreva

$$C = \frac{2\tilde{y}_2 \tilde{y}_1}{e^{\tilde{y}_2}}.$$

Começamos aplicando o Lema 2.5 (com, digamos,  $\tilde{y}_1 = \tilde{y}_2$ )

$$H_n - H_{\tilde{y}} \leq \frac{2}{\tilde{y}_{n+1}} + \frac{1}{\tilde{y}_1} \sum_{t=1}^n \frac{1}{\alpha t} \ln E e^{\tilde{y}_t(X_t - \tilde{y}_t E X_t)} \\ \leq 2 \text{ máx. } 2M \text{ em } N, \quad \frac{1}{C} \sum_{t=1}^n \frac{1}{\alpha t} \ln E e^{\tilde{y}_t(X_t - \tilde{y}_t E X_t)}.$$

Como  $\tilde{y}_t \leq 1/(2M)$ ,  $\tilde{y}_t(X_t - \tilde{y}_t E X_t) \leq 1$  e podemos aplicar a desigualdade  $e^x \leq 1 + x + (e^{\tilde{y}_2} - 2)x^2$  para todo  $x \leq 1$ . Assim, descobrimos que

$$H_n - H_{\tilde{y}} \leq 2 \text{ máx. } 2M \text{ em } N, \quad \frac{1}{C} \sum_{t=1}^n \frac{1}{\alpha t} \ln E e^{\tilde{y}_t(X_t - \tilde{y}_t E X_t)} \leq \frac{1}{C} \sum_{t=1}^n \frac{1}{\alpha t} \ln E e^{\tilde{y}_t \text{var}(X_t)}.$$

Agora denotamos por  $T$  o primeiro passo de tempo  $t$  quando  $V_t > M/2$ . Usando  $\tilde{y}_t \leq 1/(2M)$  para todos os  $t$  e  $V_T \geq M/2$ , obtemos

$$\sum_{t=1}^n \tilde{y}_t \text{var}(X_t) \leq M + \sum_{t=T+1}^n \tilde{y}_t \text{var}(X_t).$$

Limitamos a soma usando  $\tilde{y}_t \leq C \sqrt{(\ln N)/V_t \tilde{y}_1}$  para  $t \geq T$  (note que, para  $t \geq T$ ,  $V_t \tilde{y}_1 \geq V_T > M/2 > 0$ ). Isso produz

$$\sum_{t=T+1}^n \tilde{y}_t \text{var}(X_t) \leq C \sqrt{\ln N} \sum_{t=T+1}^n \frac{V_t \tilde{y}_1}{V_t \tilde{y}_1}.$$

Seja  $v_t = \text{var}(X_t) = V_t - V_t \tilde{y}_1$ . Como  $V_t \geq V_t \tilde{y}_1 + M/2$  e  $V_t \tilde{y}_1 \geq M/2$ , temos

$$\frac{v_t}{V_t \tilde{y}_1} = \frac{V_t - V_t \tilde{y}_1}{V_t \tilde{y}_1} = \frac{V_t}{V_t \tilde{y}_1} - 1 = \frac{V_t}{V_t \tilde{y}_1}.$$

Portanto,

$$\sum_{t=T+1}^n \tilde{y}_t \text{var}(X_t) \leq \frac{C \sqrt{\ln N}}{\tilde{y}_2 \tilde{y}_1} \sum_{t=T+1}^n \frac{V_t}{V_t \tilde{y}_1} = \frac{C}{\tilde{y}_2 \tilde{y}_1} \sum_{t=T+1}^n \frac{V_t}{V_t \tilde{y}_1}.$$

Substituindo nossa escolha de  $C$  e realizando superaproximações triviais, concluímos a prova. ■

**Observação 2.2.** As análises propostas pelo Teorema 2.5, Corolário 2.6 e Teorema 2.6 mostram que o predictor multilinear e o predictor de média ponderada exponencial funcionam, sem necessidade de traduzir os payoffs, tanto em jogos não assinados quanto assinados. Além disso, os limites de arrependimento mostrados nesses resultados são potencialmente muito melhores do que o limite invariante  $M \ln N$  obtido por meio da transformação explícita de payoff  $h \mapsto (h + M)/(2M)$  de jogos assinados para não assinados (consulte a Seção 2.6). No entanto, nenhum desses limites se aplica ao caso em que nenhuma informação preliminar está disponível sobre a sequência de payoffs observados.

O termo principal do limite declarado no Teorema 2.6 contém  $V_n$ . Esta quantidade é menor do que todas as quantidades da forma

$$\frac{w_{i,t} \mu_t}{\sum_{j=1}^N w_{j,t} \mu_t} h(f_{i,t}, y_t) \quad \mu_t$$

onde  $\mu_1, \mu_2, \dots$  é qualquer sequência de números reais que pode ser escolhida em *retrospecto*, pois não é necessária para a definição do predictor. Isso nos dá uma família inteira de limites superiores, e podemos escolher para a análise a sequência mais conveniente de  $\mu_t$ .

Para fornecer um exemplo concreto, denote o intervalo efetivo dos retornos no tempo  $t$  por  $R_t = \max_{i=1, \dots, N} h(f_{i,t}, y_t) - \min_{j=1, \dots, N} h(f_{j,t}, y_t)$  e considere a escolha  $\mu_t = \min_{j=1, \dots, N} h(f_{j,t}, y_t) + R_t/2$ .

**Corolário 2.7.** *Sob as mesmas suposições do Teorema 2.6,*

$$H^* - H \leq \frac{R^2 + 4M}{2} \ln N + (e - 2)M.$$

Em um jogo de perdas, onde  $h$  tem o intervalo  $[-M, 0]$ , o Corolário 2.7 afirma que o arrependimento é limitado por uma quantidade dominada pelo termo  $2M \ln N$ . Uma comparação com o limite do Teorema 2.3 mostra que perdemos apenas um fator 2 para obter uma resposta muito mais geral resultado.

## 2.9 Especialistas Simuláveis

Até agora, vimos os especialistas como entidades não especificadas gerando, a cada rodada  $t$ , um conselho ao qual o predictor tem acesso. Uma configuração diferente é quando os próprios especialistas são acessíveis ao predictor, que pode fazer experimentos arbitrários para revelar seu comportamento futuro. Neste cenário, podemos definir um especialista  $E$  usando uma sequência de funções  $f_{E,t} : Y^t \rightarrow D$ ,  $t = 1, 2, \dots$  tal que, a cada instante de tempo  $t$ , o especialista prevê de acordo com  $f_{E,t}(y^t)$ . Também assumimos que o predictor tem acesso a essas funções e, portanto, pode, a qualquer momento, hipotetizar resultados futuros e calcular todas as previsões futuras dos especialistas para essa sequência específica de resultados. Assim, o predictor pode "simular" as reações futuras dos especialistas, e chamamos esses especialistas de *simuláveis*. Por exemplo, um especialista simulável para um problema de previsão onde  $D = Y$  é o especialista  $E$  tal

que  $f_{E,t}(y_t \tilde{y}^1) = (y_1 + \dots + y_t \tilde{y}^1) / (t \tilde{y}^1)$ . Note que assumimos que no instante  $t$  a previsão de um especialista simulável depende apenas da sequência  $y_t \tilde{y}^1$  de resultados observados no passado. Isso não é verdade para o tipo mais geral de especialistas, cujo conselho pode depender de fontes arbitrárias de informação também escondidas do previsor. Mais importante, enquanto a previsão de um especialista geral, no tempo  $t$ , pode depender dos movimentos passados  $p_1, \dots, p_t \tilde{y}^1$  de o previsor (basta lembrar o protocolo do jogo de predição com conselho de especialista), os valores que um especialista simulável produz dependem apenas da sequência passada de resultados. Porque aqui não estamos preocupados com questões computacionais, permitimos que  $f_{E,t}$  seja um valor arbitrário função de  $y_t \tilde{y}^1$  e assumir que o previsor sempre pode calcular tal função.

Um tipo especial de especialista simulável é um especialista *estático*. Um especialista  $E$  é estático quando suas previsões  $f_{E,t}$  dependem apenas do índice redondo  $t$  e não de  $y_t \tilde{y}^1$ . Em outras palavras, as funções  $f_{E,t}$  são todas de valor constante. Assim, um especialista estático  $E$  é completamente descrito pela sequência  $f_{E,1}, f_{E,2}, \dots$  de suas previsões em cada rodada  $t$ . Esta sequência é fixa independentemente dos resultados observados reais. Por esta razão, usamos  $f = (f_1, f_2, \dots)$  para denotar um especialista estático arbitrário. Abusando da notação, usamos  $f$  também para denotar simulável especialistas.

Especialistas simulados e estáticos fornecem ao previsor poder adicional. É então interessante considerar se esse poder adicional poderia ser explorado para reduzir o arrependimento do previsor. Isso é investigado em profundidade para funções de perda específicas nos Capítulos 8 e 9.

## 2.10 Arrependimento Minimax

No modelo de predição com aconselhamento especializado, o melhor limite de arrependimento obtido até agora, que vale para todas as perdas convexas de valor  $[0, 1]$ , é  $\tilde{y}(n/2)$  em  $N$ . Isso é alcançado (para qualquer valor fixo  $n \tilde{y}^1$ ) pelo previsor de média ponderada exponencialmente. Esta é a melhor uniformidade possível vinculado? Que tipo de previsor atinge o melhor limite de arrependimento para cada perda específica? Para abordar essas questões de forma rigorosa, introduzimos a noção de arrependimento minimax. Corrigir uma função de perda e considere  $N$  especialistas gerais. Defina o *arrependimento minimax no horizonte*  $n$  por

$$V(N) = \min_{(f_1, 1, \dots, f_N, 1) \tilde{y}^1 D N} \max_{p^1 \tilde{y}^1 D} \min_{(f_1, 2, \dots, f_N, 2) \tilde{y}^1 D N} \max_{p^2 \tilde{y}^1 D} \min_{(f_1, n, \dots, f_N, n) \tilde{y}^1 D N} \max_{p^n \tilde{y}^1 D} \min_{t=1} \max_{e_u=1, \dots, N} \min_{t=1} (f_i, t, y_t) .$$

Uma definição equivalente, porém mais simples, de arrependimento minimax pode ser dada usando especialistas estáticos. Defina uma estratégia para o previsor como uma prescrição para calcular, em cada rodada  $t$ , a previsão  $p_t$  dados os resultados anteriores  $t \tilde{y}^1 y_1, \dots, y_t \tilde{y}^1$  e o conselho do especialista  $(f_1, s, \dots, f_N, s)$  para  $s = 1, \dots, t$ . Formalmente, uma estratégia de previsão  $P$  é uma sequência  $p_1, p_2, \dots$  de funções

$$p_t : Y_t \tilde{y}^1 \times D N \rightarrow E .$$

Agora conserte qualquer classe  $F$  de  $N$  especialistas estáticos e deixe  $L_n(P, F, y_n)$  ser a perda cumulativa no sequência  $y_n$  da estratégia de previsão  $P$  usando o conselho dos especialistas em  $F$ . Então o



arrependimento minimax  $V(N)$  pode ser definido de forma equivalente como

$$V(N) = \inf_{\mathcal{P}} \sup_{\mathcal{F}} \sup_{\mathcal{Y}} \sum_{t=1}^n \mathbb{E}_{P \in \mathcal{P}} [F_t(Y_t)] \quad \text{máx.} \quad \sum_{t=1}^n \mathbb{E}_{F \in \mathcal{F}} [L_t(P_t, Y_t)]$$

onde o ínfimo está sobre todas as estratégias de previsão  $P$  e o primeiro supremo está sobre todas as classes possíveis de  $N$  especialistas estáticos (ver Exercício 2.18).

O arrependimento minimax mede a melhor garantia de desempenho possível que se pode ter para um algoritmo de previsão que vale para todas as classes possíveis de  $N$  especialistas e todas as sequências de resultados de comprimento  $n$ . Um limite superior em  $V(N)$  estabelece a existência de uma estratégia de previsão que alcança um arrependimento não maior que o limite superior, independentemente de qual seja a classe de especialistas e a sequência de resultados. Por outro lado, um limite inferior em  $V(N)$  mostra que para qualquer estratégia de previsão existe uma classe de  $N$  especialistas e uma sequência de resultados tal que o arrependimento do preditor é pelo menos tão grande quanto o limite inferior.

Neste capítulo e no próximo, derivamos limites superiores de arrependimento minimax para várias perdas, incluindo  $V(N) \leq \ln N$  para a perda logarítmica  $(x, y) = \frac{1}{n} \sum_{t=1}^n \ln \frac{1}{p_t(y)}$  onde  $x \in [0, 1]$  e  $y \in \{0, 1\}$ , e  $V(N) \leq \frac{1}{n} \ln N$  para todas as perdas convexas de valor  $[0, 1]$ , ambas alcançadas pelo preditor de média ponderada exponencialmente. Nos Capítulos 3, 8 e 9, complementamos esses resultados provando, entre outros resultados relacionados, que  $V(N) = \ln N$  para a perda logarítmica desde que  $n \geq \log_2 N$  e que o arrependimento minimax para a perda absoluta  $(x, y) = |x - y|$  é assintoticamente  $(n/2) \ln N$ , correspondendo ao limite superior que derivamos para perdas convexas. Isso implica que o preditor de média ponderada exponencialmente é minimax ótimo, em um sentido assintótico, tanto para as perdas logarítmicas quanto para as perdas absolutas.

A noção de arrependimento minimax definida acima é baseada no desempenho de qualquer preditor no caso da pior classe possível de especialistas. No entanto, frequentemente alguém está interessado no melhor desempenho possível que um preditor pode atingir em comparação com o melhor especialista em uma classe fixa. Isso leva à definição de arrependimento minimax para uma classe fixa de especialistas (simuláveis) como segue. Fixe alguma função de perda e deixe  $F$  ser uma classe (não necessariamente finita) de especialistas simuláveis. Uma estratégia de previsão  $P$  baseada em  $F$  é agora apenas uma sequência  $p_1, p_2, \dots$  de funções  $p_t : Y \rightarrow [0, 1]$ . (Observe que  $p_t$  depende implicitamente de  $F$ , que é fixo. Portanto, como os especialistas em  $F$  são simuláveis,  $p_t$  não precisa depender explicitamente do conselho do especialista.) O *arrependimento minimax com relação a  $F$*  no horizonte  $n$  é então definido por

$$V_n(F) = \inf_P \sup_{\mathcal{F}} \sum_{t=1}^n \mathbb{E}_{P \in \mathcal{P}} [p_t(Y_t)] \quad \text{pt}(Y_t, 1), Y_t \in \mathcal{F} \quad \inf_{f \in F} \sum_{t=1}^n \mathbb{E}_{f \in F} [p_t(Y_t, 1), Y_t]$$

Essa noção de arrependimento é estudada para perdas específicas nos Capítulos 8 e 9.

Dada uma classe  $F$  de especialistas simuláveis, pode-se também definir a quantidade intimamente relacionada

$$U_n(F) = \sup_{P \in \mathcal{P}} \inf_{\mathcal{F}} \sum_{t=1}^n \mathbb{E}_{P \in \mathcal{P}} [p_t(Y_t, 1), Y_t] \quad \inf_{f \in F} \sum_{t=1}^n \mathbb{E}_{f \in F} [p_t(Y_t, 1), Y_t]$$

onde o supremo é tomado sobre todas as medidas de probabilidade sobre o conjunto  $Y_n$  de sequências de resultados de comprimento  $n$ .  $U_n(F)$  é chamado de *arrependimento maximin com relação a  $F$* . Claro, para definir medidas de probabilidade sobre  $Y_n$ , o conjunto  $Y$  de resultados deve satisfazer certas propriedades de regularidade. Para simplificar, assuma que  $Y$  é um subconjunto compacto de  $\mathbb{R}^d$ . Esta suposição é satisfeita para a maioria dos exemplos que aparecem neste livro e pode ser significativamente enfraquecida se necessário. Um teorema geral minimax, provado no Capítulo 7, implica que se a decisão

o espaço  $D$  é convexo e a função de perda é convexa e contínua em seu primeiro argumento, então

$$V_n(F) = U_n(F).$$

Esta igualdade decorre simplesmente do fato de que a função

$$F(P, Q) = \sum_{t=1}^n p_t(y_{t-1}) - \inf_{f \in F} \sum_{t=1}^n f_t(y_{t-1}) - \sum_{t=1}^n dQ(y_t)$$

é convexo em seu primeiro argumento e côncavo (na verdade linear) no segundo. Aqui definimos uma combinação convexa  $\tilde{y}P(1) + (1 - \tilde{y})P(2)$  de duas estratégias de previsão  $P(1) = (p(1)_1, p(1)_2, \dots)$  por um previsor que prevê  $\tilde{y}$  no tempo  $t$ , de acordo

para

$$\tilde{y} p(1)(y_{t-1}) + (1 - \tilde{y}) p(2)(y_{t-1}).$$

Deixamos os detalhes da verificação das condições do Teorema 7.1 para o leitor (ver Exercício 2.19).

## 2.11 Arrependimento com desconto

Em várias aplicações é razoável assumir que perdas no passado são menos significativas do que perdas sofridas recentemente. Assim, pode-se considerar *arrependimentos descontados* da forma

$$\tilde{y}_{i,n} = \sum_{t=1}^n \tilde{\gamma}_t \tilde{r}_{i,t},$$

onde os fatores de desconto  $\tilde{\gamma}_t$  são tipicamente decrescentes com  $t$  e  $r_{i,t} = (p_t, y_t) \cdot (f_{i,t}, y_t)$  é o arrependimento instantâneo com relação ao especialista  $i$  na rodada  $t$ . Em particular, assumimos que  $\tilde{\gamma}_0 \geq \tilde{\gamma}_1 \geq \tilde{\gamma}_2 \geq \dots$  é uma sequência não crescente e, sem perda de generalidade, deixamos  $\tilde{\gamma}_0 = 1$ . Assim, no tempo  $t = n$ , o arrependimento real  $r_{i,t}$  tem peso total, enquanto os arrependimentos sofridos no passado têm peso menor; quanto mais distante o passado, menor seu peso.

Nesta configuração, o objetivo do previsor é garantir que, independentemente da sequência de resultados, o arrependimento cumulativo descontado médio

$$\max_{i \in \{1, \dots, N\}} \frac{\sum_{t=1}^n \tilde{\gamma}_t \tilde{r}_{i,t}}{\sum_{t=1}^n \tilde{\gamma}_t}$$

é o menor possível. Mais precisamente, gostaríamos de limitar o arrependimento médio descontado por uma função de  $n$  que converge para zero quando  $n \rightarrow \infty$ . O propósito desta seção é explorar para quais sequências de fatores de desconto é possível atingir esse objetivo. O caso em que  $\tilde{\gamma}_t = 1$  para todo  $t$  corresponde ao caso estudado no restante deste capítulo. Outras escolhas naturais incluem a sequência de desconto exponencial  $\tilde{\gamma}_t = a^{-t}$  para algum  $a > 1$  ou sequências da forma  $\tilde{\gamma}_t = (t+1)^{-a}$  com  $a > 0$ .

Primeiro, observamos que se a sequência de desconto diminui muito rapidamente, então, exceto em casos triviais, não há esperança de provar qualquer limite significativo.

**Teorema 2.7.** *Suponha que existe uma constante positiva  $c$  tal que para cada  $n$  existem resultados  $y_1, y_2 \in Y$  e dois especialistas  $i = i(n)$  tal que  $i = \arg \min_j (f_j(n, y_1))$ ,*

$i = \arg \min_j (f(j, n, y_2))$ , e  $\min_j y_1, y_2 \mid (f(i, n, y) - \bar{y} \mid (f(i, y) - \bar{y}) \mid \bar{y} \leq c$ . Se  $\bar{y}_t < \bar{y}$ ,  $n$ ,  $\bar{y}$   
então existe uma constante  $C$  tal que, para qualquer estratégia de previsão, existe uma sequência de resultados tais que

$$\max_{eu=1, \dots, N} \frac{\sum_{t=1}^n \bar{y}_{n\bar{y}tri, t}}{\sum_{t=1}^n \bar{y}_{n\bar{y}t}} \leq C$$

para todo  $n$ .

**Prova.** O limite inferior segue simplesmente observando que o peso dos arrependimentos no último passo  $t = n$  é muito grande: é comparável ao peso total de todo o passado.

Formalmente,

$$\max_{eu=1, \dots, N} \bar{y}_{n\bar{y}tri, t} \leq \max_{eu=1, \dots, N} \bar{y}_{0ri, n} = (p_n, y_n) \leq \min_{eu=1, \dots, N} (f(i, n, y_n)).$$

Por isso,

$$\begin{aligned} e \text{ aí? } \max_{eu=1, \dots, N} \frac{\sum_{t=1}^n \bar{y}_{n\bar{y}tri, t}}{\sum_{t=1}^n \bar{y}_{n\bar{y}t}} & \leq \frac{\sup_{\bar{y}} \bar{y} (p_n, y) \leq \min_{i=1, \dots, N} (f(i, n, y))}{\sum_{t=0}^n \bar{y}_t} \\ & \leq \frac{C}{2} \cdot \blacksquare \end{aligned}$$

Em seguida, contrastamos o resultado mostrando que sempre que os fatores de desconto diminuem suficientemente lentamente, de modo que  $\bar{y}_t = \bar{y}$ , é possível fazer com que o arrependimento descontado médio desapareçam para  $n$  grande. Isso decorre de uma aplicação fácil do Teorema 2.1. Podemos definir estratégias de média ponderada com base nos arrependimentos descontados simplesmente substituindo  $r_i$  por  $r_i + \bar{y}_t = \bar{y}_{n\bar{y}tri, t}$  na definição do preditor de média ponderada. Claro, para usar tal um preditor, é preciso saber o horizonte de tempo  $n$  com antecedência. Obtemos o seguinte.

**Teorema 2.8.** Considere um preditor de média ponderada polinomial descontada definido, para  $t = 1, \dots, n$ , por

$$p_t = \frac{\sum_{j=1}^N \bar{y}_j \frac{\sum_{s=1}^{t-1} f(j, s, y_s)}{s=1+r_j, s}}{\sum_{j=1}^N \bar{y}_j \frac{\sum_{s=1}^{t-1} \bar{y}_{n\bar{y}srj, s}}{s=1+r_j, s}} = \frac{\sum_{j=1}^N \bar{y}_j \frac{\sum_{s=1}^{t-1} \bar{y}_{n\bar{y}srj, s}}{s=1+r_j, s}}{\sum_{j=1}^N \bar{y}_j \frac{\sum_{s=1}^{t-1} \bar{y}_{n\bar{y}srj, s}}{s=1+r_j, s}},$$

onde  $\bar{y}(x) = (p - \bar{y})^x p$ , com  $p = 2 \ln N$ . Então o arrependimento médio descontado satisfaz

$$\max_{eu=1, \dots, N} \frac{\sum_{t=1}^n \bar{y}_{n\bar{y}tri, t}}{\sum_{t=1}^n \bar{y}_{n\bar{y}t}} \leq 2e \ln N, \text{ não } \frac{\sum_{t=1}^n \bar{y}_{2n\bar{y}t}}{\sum_{t=1}^n \bar{y}_{n\bar{y}t}}.$$

(Um limite semelhante pode ser provado também para a previsão média ponderada exponencialmente descontada.) Em particular, se  $\bar{y}_t = \bar{y}$ , então

$$\max_{eu=1, \dots, N} \frac{\sum_{t=1}^n \bar{y}_{n\bar{y}tri, t}}{\sum_{t=1}^n \bar{y}_{n\bar{y}t}} = o(1).$$

**Prova.** Claramente o previsor satisfaz a condição de Blackwell, e para cada  $t$ ,  $|+ri,t| \leq \gamma_n \gamma_t$ .

Então o Teorema 2.1 implica, assim como na prova do Corolário 2.1,

$$\max_{i=1,\dots,N} \gamma_{i,n} \leq 2e \ln N \sum_{t=1}^n \gamma_{2t}^2$$

e a primeira afirmação segue. Para provar a segunda, basta observar que

$$\begin{aligned} \sum_{t=1}^n \frac{\gamma_{2e \ln N, n}^2}{\gamma_n \gamma_t} &= \sum_{t=1}^n \frac{\gamma_{2e \ln N, n}^2}{\gamma_n \gamma_t} = \sum_{t=1}^n \frac{\gamma_{2e \ln N, n}^2}{\gamma_n \gamma_t} \\ &= \sum_{t=1}^n \frac{\gamma_{2e \ln N, n}^2}{\gamma_n \gamma_t} = \sum_{t=1}^n \frac{\gamma_{2e \ln N, n}^2}{\gamma_n \gamma_t} \\ &= \sum_{t=1}^n \frac{\gamma_{2e \ln N, n}^2}{\gamma_n \gamma_t} = \sum_{t=1}^n \frac{\gamma_{2e \ln N, n}^2}{\gamma_n \gamma_t} = o(1). \quad \blacksquare \end{aligned}$$

É instrutivo considerar o caso especial quando  $\gamma_t = (t+1)^{-a}$  para algum  $0 < a \leq 1$ .

(Lembre-se do Teorema 2.7 que para  $a > 1$ , nenhum limite significativo pode ser derivado.) Se  $a = 1$ , o Teorema 2.8 implica que

$$\max_{i=1,\dots,N} \sum_{t=1}^n \frac{\gamma_{n \gamma_{tri}, t}}{\gamma_n \gamma_t} \leq \frac{C}{\log n}$$

para uma constante  $C > 0$ . Essa taxa lenta de convergência para zero não é surpreendente em vista do Teorema 2.7, porque a série  $1/(t+1)$  é “difícilmente não somável”. Na verdade, esse limite não pode ser melhorado substancialmente (veja Exercício 2.20). No entanto, para  $a < 1$  a convergência é mais rápida. Na verdade, um cálculo fácil mostra que o limite superior do teorema implica que

$$\max_{i=1,\dots,N} \sum_{t=1}^n \frac{\gamma_{n \gamma_{tri}, t}}{\gamma_n \gamma_t} = \begin{cases} O(1/\log n) & \text{se } a = 1 \\ O(n^a) & \text{se } 1/2 < a < 1 \\ O((\log n)/n) & \text{se } a = 1/2 \\ O(1/n) & \text{se } a < 1/2. \end{cases}$$

Não surpreendentemente, quanto mais lentamente o fator de desconto diminui, mais rápido o arrependimento cumulativo descontado médio converge para zero. No entanto, é interessante observar a “transição de fase” ocorrendo em  $a = 1/2$ : para todo  $a < 1/2$ , o arrependimento médio diminui a uma taxa  $n^{-1/2}$ , um comportamento quantitativamente similar ao caso quando nenhum desconto é levado em conta.

## 2.12 Observações bibliográficas

Nosso modelo de predição sequencial com aconselhamento especializado encontra suas raízes na teoria de jogos repetidos. Jogos repetidos de soma zero com matriz de perdas fixas são um tópico clássico da teoria dos jogos. Nesses jogos, o arrependimento após  $n$  jogadas é definido como a perda excessiva do jogador da linha em relação à menor perda que poderia ser incorrida se ele conhecesse com antecedência a distribuição empírica das ações do jogador da coluna durante as  $n$  jogadas. Em seu trabalho pioneiro, Hannan [141] cria uma estratégia de jogo aleatória cuja rodada por

o arrependimento esperado cresce a uma taxa  $N\sqrt{3nm}/2$ , onde  $N$  é o número de linhas,  $m$  é o número de colunas na matriz de perdas e  $n$  é o número de jogadas. Conforme mostrado no Capítulo 7, nossos previsores de média ponderada polinomial e exponencial podem ser usados para jogar jogos repetidos de soma zero, alcançando o arrependimento  $\sqrt{n}/2$  em  $N$ . Obtemos a mesma dependência do número  $n$  de jogadas em comparação com o arrependimento de Hannan, mas melhoramos significativamente a dependência das dimensões,  $N$  e  $m$ , da matriz de perdas. Um jogador randomizado diferente com um arrependimento por rodada que desaparece também pode ser derivado do famoso teorema de acessibilidade de Blackwell [28], generalizando o teorema minimax de von Neumann para recompensas com valor vetorial. Este resultado, que derivamos novamente no Capítulo 7, é baseado em uma estratégia mista equivalente ao nosso predictor de média ponderada polinomial com  $p = 2$ . Constantes assintóticas exatas para o arrependimento minimax (para um caso especial) foram mostradas pela primeira vez por Cover [68]. Em nossa terminologia, Cover investiga o problema de prever uma sequência de resultados binários com dois especialistas estáticos, um sempre prevendo 0 e o outro sempre prevendo 1. Ele mostra que o arrependimento minimax para a perda absoluta neste caso especial é  $(1 + o(1))\sqrt{n}/(2\sqrt{2})$ .

O problema da predição sequencial, desprovido de qualquer suposição probabilística, está profundamente conectado com o problema da teoria da informação de comprimir uma sequência de dados individual. Uma pesquisa pioneira neste campo foi realizada por Ziv [317, 318] e Lempel e Ziv [197, 317, 319], que resolveram o problema de comprimir uma sequência de dados individual quase tão bem quanto o melhor autômato de estado finito. Conforme mostrado por Feder, Merhav e Gutman [95], o compressor Lempel-Ziv pode ser usado como um predictor randomizado (para a perda absoluta) com um arrependimento por rodada que desaparece contra a classe de todos os especialistas em estado finito, um resultado surpreendente considerando a rica estrutura desta classe. Além disso, Feder, Merhav e Gutman elaboram, para a mesma classe de especialistas, um predictor com uma taxa de convergência melhor do que a taxa demonstrável para o predictor Lempel-Ziv (veja também Merhav e Feder [213] para mais resultados ao longo dessas linhas). Na Seção 9, continuamos a investigação da relação entre predição e compressão, mostrando condições simples sob as quais a predição com perda logarítmica é minimax equivalente à compressão adaptativa de dados. As conexões entre predição com aconselhamento especializado e conteúdo de informação de uma sequência individual foram exploradas por Vovk e Watkins [303], que introduziram a noção de complexidade preditiva de uma sequência de dados, uma quantidade que, para a perda logarítmica, está relacionada à complexidade de Kolmogorov da sequência. Referimo-nos ao livro de Li e Vitanyi [198] para uma excelente introdução à teoria algorítmica da aleatoriedade.

Aproximadamente na mesma época em que Hannan e Blackwell estavam estabelecendo as bases da abordagem da teoria dos jogos para a previsão, Solomonoff teve a ideia de formalizar o fenômeno da inferência indutiva em humanos como um processo de previsão sequencial. Essa pesquisa eventualmente o levou à introdução de uma *probabilidade universal prior* [273–275], a ser usada como prior na inferência bayesiana. Um importante “produto secundário” do prior universal de Solomonoff é a noção de aleatoriedade algorítmica, que ele introduziu independentemente de Kolmogorov. Embora reconheçamos o papel fundamental desempenhado por Solomonoff no campo da teoria da previsão sequencial, especialmente em conexão com a complexidade de Kolmogorov, neste livro olhamos para o problema da previsão de um ângulo diferente. Dito isso, certamente achamos que explorar as conexões entre aleatoriedade algorítmica e teoria dos jogos, por meio da noção unificadora de previsão, é um plano de pesquisa certamente empolgante.

O campo da inferência indutiva investiga o problema da previsão sequencial quando os especialistas são funções retiradas de uma grande classe, possivelmente incluindo todas as linguagens recursivas ou todas as funções recursivas parciais, e a tarefa é eventualmente identificar um especialista que

é consistente (ou quase consistente) com uma sequência infinita de observações. Esta aprendizagem paradigma, introduzido em 1967 por Gold [130], ainda é ativamente estudado. Ao contrário da teoria descrito neste livro, cujas raízes são a teoria dos jogos, as principais ideias e ferramentas analíticas usado na inferência indutiva vem da teoria da recursão (ver Odifreddi [227]).

Na ciência da computação, uma área relacionada à previsão com especialistas é a análise competitiva de algoritmos online (veja a monografia de Borodin e El-Yaniv [36] para uma pesquisa). Um bom exemplo de um artigo explorando o uso de algoritmos de previsão em análise competitiva é o trabalho de Blum e Burch [32].

O paradigma da previsão com aconselhamento especializado foi introduzido por Littlestone e Warmuth [203] e Vovk [297], e posteriormente desenvolvido por Cesa-Bianchi, Freund, Haussler, Helmbold, Schapire e Warmuth [48] e Vovk [298], embora alguns dos seus principais ingredientes já aparecem nos artigos de De Santis, Markowski e Wegman [260], Littlestone [200] e Foster [103]. O uso de funções potenciais na predição sequencial é devido a Hart e Mas-Colell [146], que usaram a condição de Blackwell em uma teoria de jogos contexto, e para Grove, Littlestone e Schuurmans [133], que usaram exatamente o mesmo condição para a análise de certas variantes do algoritmo Perceptron (ver Capítulo 11).

Nosso Teorema 2.1 é inspirado e parcialmente baseado na análise de Hart e Mas-Colell de

-estratégias para jogar jogos repetidos [146] e na análise do algoritmo quase-aditivo de Grove, Littlestone e Schuurmans [133]. A estrutura unificada para sequencial

A previsão baseada em funções potenciais que descrevemos aqui foi introduzida por Cesa-Bianchi e Lugosi [54]. A previsão baseada no potencial exponencial tem sido usada em

teoria dos jogos como uma variante do jogo fictício suave (ver, por exemplo, o livro de Fudenberg e Levine [119]). Na teoria da aprendizagem, os previsores de média ponderada exponencial foram introduzidos e analisados por Littlestone e Warmuth [203] (o algoritmo da maioria ponderada)

e por Vovk [297] (o algoritmo de agregação). O truque de definir o parâmetro  $p$  de o potencial polinomial para  $2 \ln N$  é devido a Gentile [123]. A análise na Seção 2.2 é

com base no trabalho de Cesa-Bianchi [46]. A ideia do truque de duplicação da Seção 2.3 aparece em os artigos de Cesa-Bianchi, Freund, Haussler, Helmbold, Schapire e Warmuth [48] e

Vovk [298], enquanto a análise do Teorema 2.3 é adaptada de Auer, Cesa-Bianchi,

e Gentile [13]. Os limites dependentes de dados da Seção 2.4 são baseados em duas fontes:

O Teorema 2.4 é do trabalho de Littlestone e Warmuth [203] e o Corolário 2.4 é devido

para Freund e Schapire [112]. Uma análise mais sofisticada da ponderação exponencial

o previsor médio com  $\gamma$  variável no tempo é devido a Yaroshinski, El-Yaniv e Seiden [315].

Eles mostram um limite de arrependimento da ordem  $(1 + o(1)) 2L\gamma$  Em  $N$ , onde  $o(1) \rightarrow 0$  para  $L\gamma \rightarrow \infty$ .

Hutter e Poland [165] provam um resultado semelhante ao Exercício 2.10 usando follow-the-perturbed-leader, um previsor aleatório que analisamos no Capítulo 4.

O previsor multilinear e os resultados da Seção 2.8 são devidos a Cesa-Bianchi, Man-sour e Stoltz [57].

Uma versão mais fraca do Corolário 2.7 foi provada por Allenberg-Neeman

e Neeman [7].

O previsor baseado em gradiente da Seção 2.5 foi introduzido por Kivinen e Warmuth [181]. A prova do Corolário 2.5 é devida a Cesa-Bianchi [46]. A noção de especialistas simuláveis e arrependimento do pior caso para a estrutura dos especialistas foi investigada pela primeira vez por

Cesa-Bianchi et al. [48]. Os resultados para funções de perda mais gerais estão contidos no Chung's

artigo [60]. Fudenberg e Levine [121] consideram que arrependimentos descontados são uma situação um tanto diferente modelo diferente do discutido aqui.

O modelo de predição com aconselhamento especializado está conectado à teoria da decisão bayesiana. Para por exemplo, quando a perda absoluta é usada, os pesos normalizados da média ponderada

previsor baseado no potencial exponencial aproxima bastante a distribuição posterior de um modelo generativo estocástico simples para a sequência de dados (veja Exercício 2.7). Deste ponto de vista, nossa análise de arrependimento mostra um exemplo em que as decisões de Bayes são robustas em um sentido forte, porque seu desempenho pode ser limitado não apenas em expectativa com relação ao sorteio aleatório da sequência, mas também para cada sequência individual.

## 2.13 Exercícios

**2.1** Suponha que você tenha que prever uma sequência  $Y_1, Y_2, \dots \in \{0, 1\}$  de variáveis aleatórias iid com distribuição desconhecida, seu espaço de decisão é  $[0, 1]$  e a função de perda é  $\ell(p, y) = |p - y|$ .

Como você procederia? Tente estimar a perda cumulativa do seu previsor e compare-a com a perda cumulativa do melhor dos dois especialistas, um dos quais sempre prevê 1 e o outro sempre prevê 0. Quais são as distribuições mais "difíceis"? Como seu arrependimento (esperado) se compara ao do algoritmo de média ponderada (que não "sabe" que a sequência de resultados é iid)?

**2.2** Considere um previsor de média ponderada com base em uma função potencial

$$\phi(u) = \sum_{i=1}^N \phi(u_i).$$

Suponha ainda que a quantidade  $C(\mathbf{r}_t)$  que aparece na declaração do Teorema 2.1 é limitada por uma constante para todos os valores de  $\mathbf{r}_t$  e que a função  $\phi(\mathbf{y}(u))$  é estritamente convexa. Mostre que existe uma sequência não negativa  $\tilde{y}_n \geq 0$  tal que o arrependimento cumulativo do previsor satisfaz, para todo  $n$  e para toda sequência de resultados  $y_n$ ,

$$\sum_{i=1}^n R_{i,n} \leq \tilde{y}_n.$$

**2.3** Analise o previsor de média ponderada polinomial usando o Teorema 2.1, mas usando a função potencial  $\phi(u) = u + p$  em vez da escolha  $\phi(u) = u$  usada na prova do Corolário 2.1. Derive um limite da mesma forma que no Corolário 2.1, talvez com constantes diferentes.

**2.4** Seja  $Y = \{0, 1\}$ ,  $D = [0, 1]$ , e  $\ell(p, y) = |p - y|$ . Prove que a perda cumulativa  $L$  do previsor de média ponderada exponencialmente é sempre pelo menos tão grande quanto a perda cumulativa  $\min_j L_j$  do melhor especialista. Mostre que para outras funções de perda, como a perda quadrada  $(p - y)^2$ , isso não é necessariamente assim. Dica: Tente reverter a prova do Teorema 2.2.

**2.5 (Pesos iniciais não uniformes)** Por definição, o previsor de média ponderada usa  $N$  uniforme. No entanto, não pesos iniciais  $w_{i,0} = 1$  para todo  $i = 1, \dots, N$ . escolha, há nada de especial nisso e a análise do arrependimento para este previsor pode ser realizada usando qualquer conjunto de números não negativos para os pesos iniciais.

Considere a execução do previsor de média ponderada exponencialmente com pesos iniciais arbitrários  $w_{1,0}, \dots, w_{N,0} > 0$ , definido, para todo  $t = 1, 2, \dots$ , por

$$p_t = \frac{\sum_{i=1}^N w_{i,t} \tilde{y}_{i,t}}{\sum_{j=1}^N w_{j,t}}, \quad w_{i,t} = w_{i,0} e^{-\sum_{j=1}^t \ell(p_{j-1}, y_j)} \quad (f_{i,t}, y_t).$$

Nas mesmas condições do enunciado do Teorema 2.2, mostre que para cada  $n$  e para cada sequência de resultados  $y_n$ ,

$$L_n \leq \min_{i=1, \dots, N} L_{i,n} + \ln \tilde{y}_n \leq \frac{1}{W_0} \sum_{i=1}^N \frac{1}{w_{i,0}} \ln \frac{1}{w_{i,0}} + \frac{\tilde{y}_n}{8},$$

onde  $W_0 = w_{1,0} + \dots + w_{N,0}$ .

**2.6 (Muitos bons especialistas)** Sequências de resultados em que muitos especialistas sofrem uma pequena perda são intuitivamente mais fácil de prever. Adapte a prova do Teorema 2.2 para mostrar que o exponencialmente o predictor ponderado satisfaz a seguinte propriedade: para cada  $n$ , para cada sequência de resultados  $y_1, \dots, y_n$ , e para todo  $L > 0$ ,

$$L_n \leq L + \frac{1}{\gamma} \sum_{i=1}^n \frac{L_i}{L} + \frac{\gamma}{8}$$

onde  $L$  é a cardinalidade do conjunto  $\{1, \dots, N\}$ .

**2.7 (Geração aleatória da sequência de resultados)** Considere a média ponderada exponencialmente predictor e definir o seguinte modelo probabilístico para a geração da sequência  $y_1, \dots, y_n$ , onde agora vemos cada bit  $y_t$  como a realização de uma variável aleatória de Bernoulli  $Y_t$ . Um especialista  $i$  é escolhido aleatoriamente do conjunto de  $N$  especialistas. Para cada  $t = 1, \dots, n$ ,  $X_t \in \{0, 1\}$  é desenhado de modo que  $X_t = 1$  com probabilidade  $f_t$ . Então  $Y_t$  é definido como  $X_t$  com probabilidade  $\gamma$  e é definido como  $1 - X_t$  com probabilidade  $1 - \gamma$ , onde  $\gamma = 1/(1 + e^{\gamma})$ . Mostre que o predictor pesos  $w_i(t)/(w_1(t) + \dots + w_N(t))$  e são iguais à probabilidade posterior  $P[i = i | Y_1 = y_1, \dots, Y_t = y_t]$  que o especialista  $i$  é desenhado dado que a sequência  $y_1, \dots, y_t$  foi observado.

**2.8 (O truque da duplicação)** Considere a seguinte estratégia de previsão ("truque da duplicação"): o tempo é dividido em períodos  $(2^m, \dots, 2^{m+1} - 1)$ , onde  $m = 0, 1, 2, \dots$ . No período  $(2^m, \dots, 2^{m+1} - 1)$ , a estratégia usa o predictor de média ponderada exponencialmente inicializado no tempo  $2^m$  com parâmetro  $\gamma_m = \gamma/2^m$ . Assim, o predictor de média ponderada é reiniciado a cada momento instância que é uma potência inteira de 2 e é reiniciada com um novo valor de  $\gamma$ . Usando o Teorema 2.2 provar que, para qualquer sequência  $y_1, y_2, \dots, y_n$  de resultados e para qualquer  $n \geq 1$ , o arrependimento desta o meteorologista é no máximo

$$L_n \leq \min_{i=1, \dots, N} L_i + \frac{\gamma}{2} \sum_{t=1}^n \frac{L_t}{L} + \frac{\gamma}{2} \log N.$$

**2.9 (O truque da duplicação, continuação)** No Exercício 2.8, de forma bastante arbitrária, dividimos o tempo em períodos de comprimento  $2^m$ ,  $m = 1, 2, \dots$ . Investigue o que acontece se, em vez disso, os comprimentos dos períodos forem de forma  $a^m$  para algum outro valor de  $a > 0$ . Qual escolha de  $a$  minimiza, assintoticamente, o constante no limite? Quanto você pode ganhar em comparação com o limite dado no texto?

**2.10** Combine o Teorema 2.4 com o truque de duplicação do Exercício 2.8 para construir um predictor que, sem qualquer conhecimento prévio de  $L$ , alcança, para todo  $n$ ,

$$L_n \leq \gamma \sum_{i=1}^n L_i + \frac{2}{\gamma} \log N + c$$

sempre que a função de perda for limitada e convexa em seu primeiro argumento, e onde  $c$  é uma constante positiva.

**2.11 (Outro potencial de variação temporal)** Considere a média ponderada exponencial adaptativa predictor que, no tempo  $t$ , usa

$$y_t = c \frac{e^{-\gamma \sum_{i=1}^t L_i}}{\sum_{i=1}^N e^{-\gamma \sum_{i=1}^t L_i}},$$

onde  $c$  é uma constante positiva. Mostre que sempre que uma função de perda com valor  $[0, 1]$  é convexa em seu primeiro argumento, então existe uma escolha de  $c$  tal que

$$L_n \leq \gamma \sum_{i=1}^n L_i + \frac{2}{\gamma} \log N + \gamma$$

onde  $\gamma > 0$  é uma constante apropriada (Auer, Cesa-Bianchi e Gentile [13]). Dica: Siga o esboço da prova do Teorema 2.3. Este exercício não é fácil.

**2.12** Considere o problema de previsão com  $Y = D = [0, 1]$  com a perda absoluta  $\ell(p, y) = |p - y|$ . Mostre que neste caso o predictor de média ponderada exponencialmente baseado em gradiente coincide



com o predictor de média ponderada exponencialmente. (Observe que a derivada da perda não existe para  $p = y$  e a definição da média ponderada exponencialmente baseada em gradiente O meteorologista precisa ser ajustado adequadamente.)

**2.13** Prove o Lema 2.4.

**2.14** Use o truque da duplicação para provar uma variante do Corolário 2.6 em que nenhum conhecimento sobre a sequência de resultados é assumida como preliminarmente disponível (no entanto, como no corolário, ainda assumimos que a função de recompensa  $h$  tem alcance  $[yM, M]$  e é côncava em seu primeiro argumento). Expresse o limite de arrependimento em termos do menor limite superior monótono no sequência  $Q_1, Q_2, \dots$  (ver Cesa-Bianchi, Mansour e Stoltz [57]).

**2.15** Prove uma variante do Teorema 2.6 em que nenhum conhecimento sobre o intervalo  $[yM, M]$  do payoff supõe-se que a função esteja preliminarmente disponível (ver Cesa-Bianchi, Mansour e Stoltz [57]).  
*Dica:* Substitua o termo  $1/(2M)$  na definição de  $\tilde{y}_t$  por  $2\tilde{y}_t(1+kt)$ , onde  $k$  é o menor inteiro não negativo tal que  $\max_{s=1,\dots,t} \tilde{y}_s \max_{i=1,\dots,N} |h(f_{i,s}, y_s)| \leq 2k$ .

**2.16** Prove um limite de arrependimento para o predictor multilinear usando a atualização  $w_{i,t} = w_{i,t-1}(1 + \tilde{y}_{ri,t})$ , onde  $r_{i,t} = h(f_{i,t}, y_t) - h(p_t, y_t)$  é o arrependimento instantâneo. O que você pode dizer sobre o evolução do peso total  $W_t = w_{1,t} + \dots + w_{N,t}$  dos especialistas?

**2.17** Prove o Lema 2.5. *Dica:* Adapte a prova do Teorema 2.3.

**2.18** Mostre que as duas expressões do arrependimento minimax  $V(N)$  na Seção 2.10 são equivalentes.

**2.19** Considere uma classe  $F$  de especialistas simuláveis. Suponha que o conjunto  $Y$  de resultados seja um conjunto compacto subconjunto de  $\mathbb{R}^d$ , o espaço de decisão  $D$  é convexo e a função de perda é convexa e contínua em seu primeiro argumento. Mostre que  $V_n(F) = U_n(F)$ . *Dica:* Verifique as condições do Teorema 7.1.

**2.20** Considere os fatores de desconto  $\tilde{y}_t = 1/(t+1)$  e suponha que haja uma constante positiva  $c$  de modo que para cada  $n$  existam resultados  $y_1, y_2 \in Y$  e dois especialistas  $i = i$  tais que  $i = \arg\min_j (f_{j,n}, y_1)$ ,  $i = \arg\min_j (f_{j,n}, y_2)$ , e  $\min_j y_1, y_2 \mid (f_{i,n}, y) \leq (f_{i,n}, y) \leq c$ . Mostrar que existe uma constante  $C$  tal que para qualquer estratégia de previsão, existe uma sequência de resultados tais que

$$\max_{u=1,\dots,N} \frac{\sum_{t=1}^n \tilde{y}_t \tilde{y}_{tri,t}}{\sum_{t=1}^n \tilde{y}_t} \leq \frac{C}{\text{registro } n}$$

para todo  $n$ .