Exercícios sobre Modelos Lineares

No que segue, sempre suponha que as variáveis aleatória relevantes estão definidas no mesmo espaço de probabilidade.

**Exercício 1** Seja $Z$ uma variável aleatória discreta, que toma valores no conjunto $\{z_1, \ldots, z_n\}$; e $Y$ uma variável aleatória com segundo momento finito. Defina o vetor aleatório $X = (\mathbf{1}_{\{z_2\}}(Z), \ldots \mathbf{1}_{\{z_n\}}(Z))$. Mostre que o melhor preditor linear de $Y$ como função de um intercepto e $X$ é da forma:

$$\alpha_* + \delta_*' X$$

com

$$\alpha_* = \mathbb{E}[Y|Z = z_1]$$
$$\delta_j = \mathbb{E}[Y|Z = z_{j+1}] - \mathbb{E}[Y|Z = z_1]$$

onde definimos o objeto $\mathbb{E}[Y|Z = z_j]$ como:

$$\mathbb{E}[Y|Z = z_j] := \frac{\mathbb{E}[Y \mathbf{1}_{\{z_j\}}(Z)]}{\mathbb{P}[\{Z = z_j\}]} .$$

. Conclua que o melhor preditor linear coincide com $\mathbb{E}[Y|Z]$.

**Exercício 2** Seja $h^*$ a função de esperança condicional de uma variável aleatória $Y$ com segundo momento finito dado uma variável aleatória real $X$, i.e. $h^*(X) = \mathbb{E}[Y|X]$. Compute o melhor preditor linear de $Y$ como função de um intercepto e $X$, para os casos abaixo. Ilustre graficamente a derivada de $h^*$ comparando-a com a inclinação do melhor preditor linear, nos pontos do suporte[1] de $X$. Identifique as regiões onde o melhor preditor linear mais "erra" a inclinação de $h^*$.

a $h^*(x) = x^2$ e $X \sim U[0, 1]$.

b $h^*(x) = \min\{x^3, 100\}$ e $X \sim$ Exponecial(1).

c $h^*(x) = x^2$ e $X \sim \chi^2(1)$.

**Exercício 3** Seja $A$ uma matriz simétrica positiva semidefinida:

a Mostre que os autovalores de $A$ sempre são não negativos. Dica: Se $v \neq 0$ é um autovetor de um autovalor $\lambda$, $Av = \lambda v$.

b Mostre que o menor autovalor de $A$ é zero se, e somente se, existe $v \neq 0$, $v'Av = 0$. Dica: para suficiência, utilize a decomposição espectral de uma matriz simétrica.

[1] O suporte de uma variável aleatória $X$ é definido como o menor conjunto fechado $C$ tal que $\mathbb{P}[X \in C] = 1$.

**Exercício 4** Seja $(\Omega, \Sigma, \mathbb{P})$ um espaço de probabilidade, em que cada elemento $\omega \in \Omega$ corresponde a uma firma num setor de interesse, e $\mathbb{P}[A]$ denota a proporção de firmas do "tipo" A . Considere o exemplo, visto em aula, em que a função de produção no setor é $h(z, l) = zl^\alpha$, $\alpha < 1$. Suponha que, no momento da decisão de produção, uma firma $\omega \in \Omega$ na população somente observe um sinal $\xi(\omega)$ de sua produtividade. A distribuição do sinal $\xi$ na população é denotada por $\mathbb{P}_\xi$. As firmas então escolhem $l$ otimizando:

$$l(\omega) \in \text{argmax}_{l \geq 0}\, \xi(\omega) l^\alpha - wl\,,$$

onde $w$ denota o salário (em unidades do bem produzido pelo setor), aqui tratado como constante fixa.[2] Após a escolha de demanda por trabalho, a produtividade de cada firma é revelada como $z(\omega) = \xi(\omega)\epsilon(\omega)$, onde o "choque" ou "surpresa" na produtividade não é antecipável com base no sinal, i.e. $\xi$ é independente de $\epsilon$ ($\mathbb{P}_{(\xi,\epsilon)} = \mathbb{P}_\xi \otimes \mathbb{P}_\epsilon$). A produção final das firmas é dada por $Y(\omega) = z(\omega)l(\omega)$.

a Encontre o melhor preditor linear da variável aleatória $\log(Y)$ como função de um intercepto e $\log(l)$. O coeficiente associado a $\log(l)$ coincide com $\alpha$? Por quê?

b Suponha que, caso $Y(\omega) - wl(\omega) < 0$, a firma $\omega$ não opera, e, nesse caso, a demanda por trabalho e quantidade produzida observadas são zero. Calcule o melhor preditor linear nesse caso. Ele corresponde ao obtido anteriormente? Por quê?

[2] Em outras palavras, a incerteza capturada pelas variáveis aleatórias só se deve à amostragem no sentido clássico, e não à incerteza econômica.