Lista I - MAE0328

Guilherme Navarro NºUSP: 8943160

Questão 1

Considere a matriz $A = \begin{pmatrix} \alpha_1 & \beta \\ \beta & \alpha_2 \end{pmatrix}$ com $\alpha_1, \alpha_2, \beta \in \mathbb{R}$. Prove a seguinte proposição: *Se* $det(A) > 0$, *e* $\alpha_1, \alpha_2 > 0$ *então* $x^T A x > 0 \ \forall \ x \in \mathbb{R}, \ x \neq (0,0)^T$.

Resolução

$det(A) = \alpha_1\alpha_2 - \beta^2 > 0 \Rightarrow \beta^2 < \alpha_1\alpha_2 \Rightarrow \beta < \sqrt{\alpha_1\alpha_2}$ (I) com $\alpha_1, \alpha_2 > 0$

Seja $x^T = (x_1 \ x_2)$ e $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ Mostrar que:

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \alpha_1 & \beta \\ \beta & \alpha_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} > 0$$

Então,

$$\begin{pmatrix} \alpha_1 x_1 + \beta x_2 & \beta x_1 + \alpha_2 x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \alpha_1 x_1^2 + \beta x_1 x_2 + \beta x_1 x_2 + \alpha_2 x_2^2 = \alpha_1 x_1^2 + 2\beta x_1 x_2 + \alpha_2 x_2^2 \ (II)$$

Sabemos que: $\alpha_1 x_1^2 > 0$ e $\alpha_2 x_2^2 > 0$ substituindo β da equação (II) pelo da equação (I), temos que:

$$\alpha_1 x_1^2 + 2\sqrt{\alpha_1\alpha_2} x_1 x_2 + \alpha_2 x_2^2 \ (III)$$

Assim, podemos escrever (III) da seguinte forma:

$$(\sqrt{\alpha_1} x_1 + \sqrt{\alpha_2} x_2)^2 > 0 \Rightarrow x^T A x > 0 \ \blacksquare$$

Questão 2

Os próximos itens são referentes ao método iterativo de Newton-Raphson para $\theta = (\mu, \sigma)$,

$$\begin{pmatrix} \hat{\mu}^{(j+1)} \\ \hat{\sigma}^{(j+1)} \end{pmatrix} = \begin{pmatrix} \hat{\mu}^{(j)} \\ \hat{\sigma}(j) \end{pmatrix} - \begin{pmatrix} \frac{\partial^2 l(\theta,y)}{\partial \mu^2} & \frac{\partial^2 l(\theta,y)}{\partial \mu \partial \sigma} \\ \frac{\partial^2 l(\theta,y)}{\partial \mu \partial \sigma} & \frac{\partial^2 l(\theta,y)}{\partial \sigma^2} \end{pmatrix}^{-1}_{(\theta = \hat{\theta}^{(j)})} \begin{pmatrix} \frac{\partial l(\theta,y)}{\partial \mu} \\ \frac{\partial l(\theta,y)}{\partial \sigma} \end{pmatrix}_{(\theta = \hat{\theta}^{(j)})}, \ j = \in \{1, ..., n\}$$

(a) Considere o modelo $Y = \mu + e_i$, em que $w_i = \frac{1}{\sigma} e_i \overset{iid}{\sim} \ Cauchy(0, 1), \ i \in \{1, ..., n\}$ i.e.

$$f_w(w_i) = \frac{1}{\pi(1 + w_i^2)} \mathbb{I}_R(w_i)$$

Calcule o logaritmo da função de verossimilhança $l(\theta, y)$ e suas primeiras e segundas derivadas.

Resolução

Aplicando o teorema da transformação de variável: $Y = \mu + \sigma w_i \Rightarrow Y^{-1} = \frac{y_i - \mu}{\sigma} \Rightarrow \frac{dY^{-1}}{dy_i} = \frac{1}{\sigma}$ Logo:

$$f_y(\frac{y_i - \mu}{\sigma})|\frac{1}{\sigma}| = f_y(\frac{y_i - \mu}{\sigma})\frac{1}{\sigma} = \frac{1}{\pi\sigma(1 + (\frac{y_i - \mu}{\sigma})^2)} \mathbb{I}_R(y_i)$$

Assim $Y_i \sim Cauchy(\mu, \sigma) \Rightarrow f(y_i) = \frac{1}{\pi\sigma[1 + (\frac{y_i - \mu}{\sigma})^2]} = \frac{1}{\pi[\sigma + \frac{(y_i - \mu)^2}{\sigma}]} = \frac{\sigma}{\pi(\sigma^2 + (y_i - \mu)^2)} \mathbb{I}_R(y_i) \ i \in \{1, ..., n\}$ com $\sigma > 0$. Logo a sua função de verossimilhança é:

$$L(\theta, y) = \prod_{i=1}^{n} f(y_i) = \prod_{i=1}^{n} \frac{\sigma}{\pi(\sigma^2 + (y_i - \mu)^2)} = (\frac{\sigma}{\pi})^n \prod_{i=1}^{n} \frac{1}{\sigma^2 + (y_i - \mu)^2}$$

Aplicando o ln(logaritmo neperiano) em $L(\theta, y)$. Seja $l(\theta, y) = ln(L(\theta, y))$

$$l(\theta, y) = n * ln(\sigma) - n * ln(\pi) - \sum_{i=1}^{n} ln(\sigma^2 + (y_i - \mu)^2)$$

Calculando suas primeiras derivadas:

$$\frac{\partial l(\theta, y)}{\partial \theta} = n \begin{pmatrix} 0 \\ \frac{1}{\sigma} \end{pmatrix} - n * ln(\pi) \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \sum_{i=1}^{n} \begin{pmatrix} \frac{-2(y_i - \mu)}{\sigma^2 + (y_i - \mu)^2} \\ \frac{2\sigma}{\sigma^2 + (y_i - \mu)^2} \end{pmatrix}$$

Agora, calculando suas segundas derivadas:

$$\frac{\partial^2 l(\theta, y)}{\partial \theta^T \partial \theta} = n \begin{pmatrix} 0 & 0 \\ 0 & -\frac{1}{\sigma^2} \end{pmatrix} - n * ln(\pi) \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} - \sum_{i=1}^{n} \begin{pmatrix} \frac{-2(y_i - \mu)^2 - \sigma^2}{(\sigma^2 + (y_i - \mu)^2)^2} & \frac{-4\sigma(y_i - \mu)}{(\sigma^2 + (y_i - \mu)^2)^2} \\ \frac{-4\sigma(y_i - \mu)}{(\sigma^2 + (y_i - \mu)^2)^2} & \frac{-(y_i - \mu)^2 - \sigma^2}{(\sigma^2 + (y_i - \mu)^2)^2} \end{pmatrix}$$

(b) Implemente no R o método de Newton-Raphson usando o exercício anterior. Proponha estimativas iniciais para iniciar o processo iterativo e considere como critério de parada do processo iterativo um erro máximo de 10^{-6}.

Resolução

Propondo uma estimativa inicial $\begin{pmatrix} \hat{\mu}^{(0)} \\ \hat{\sigma}^{(0)} \end{pmatrix} = \begin{pmatrix} mediana(dados) \\ 1 \end{pmatrix} = \begin{pmatrix} 15.1 \\ 1 \end{pmatrix}$

Obs: Adaptação do algoritmo Newton-Raphson feito pelo professor Patriota na disciplina de Inferência (2^{o}/2018)

```r
library(MASS)

dados <- c(17.33,-90.91,18.49,25.03,19.88,18.51,12.39,11.24,11.78,15.15,
       12.24,-238.41,15.19,24.51,11.82,18.16,15.73,11.26,7.41,19.46,
       25.58,-22.27,16.72,17.37,17.42,13.37,12.15,13.75,24.93,34.49,
       21.66,-7.00,16.21,42.07,-2.70,14.91,16.36,14.57,15.11,12.63,
       2.96,13.24,18.68,19.22,30.43,12.87,11.40,-68.39,14.43,13.28,
       3.24,12.63,11.92,22.08,19.47,13.84,20.68,-10.52,12.78,19.03,
       16.65,16.46,2.65,13.55,15.61,59.20,11.46,17.26,18.17,21.50,
       20.31,9.63,13.84,17.94,16.35,18.40,-16.15,17.27,255.83,20.74,
       -1.08,13.67,-5.82,-0.75,21.05,16.70,23.67,11.82,6.23,17.61,
       5.83,19.83,11.14,18.06,12.11,19.83,16.20,11.03,13.99,8.70)

MV = function(x, theta0, epsilon=10^(-6), it=1000){ # Dados, esti. inicial, erro, iterações
  n <- length(x)

  U <- function(t1,t2){ # Primeiras derivadas
    U1 <- 2*sum((x-t1)/(t2^2+(x-t1)^2))
    U2 <- n/t2 - 2*sum(t2/(t2^2 + (x-t1)^2))
    return (c(U1,U2))
  }

  H <- function(t1,t2){ # Segundas derivadas
    H11 <- 2*sum(((x-t1)^2-t2^2)/((x-t1)^2+t2^2)^2)
    H12 <- 4*t2*sum((x-t1)/((x-t1)^2+t2^2)^2)
    H22 <- sum(((x-t1)^2-t2^2)/((x-t1)^2+t2^2)^2) - n/t2^2
    M <- matrix(c(H11,H12,H12,H22),2,2)
    return (M)
  }

  erro <- 10
  j <- 0
  t1 <- numeric () # mi
  t2 <- numeric() # Sigma
  t1[1]<-theta0[1] # recebe a estimativa inicial
  t2[1]<-theta0[2] # recebe a estimativa inicial

  while(erro > epsilon & j < it){
    j <- j+1
    Aux <- c(t1[j],t2[j]) - solve(H(t1[j],t2[j]))%*%U(t1[j],t2[j])
    t1[j+1] <- Aux[1]
    t2[j+1] <- Aux[2]
    erro <- max(abs(c(t1[j+1]-t1[j],t2[j+1]-t2[j])))
  }

  S <- list()
  S$erro <- erro
  S$Iteracoes <- j
  S$theta <- c(t1[j+1],t2[j+1])
  S$H <- H(t1[j+1],t2[j+1])
  S$U <- U(t1[j+1],t2[j+1])
  return(S)
}
```

```r
#Encontrando as estimativas de MV para o exercício
theta0 <- c(median(dados),1) # Estimativa inicial
fit <- MV(dados,theta0)

# Estimativas:
fit$theta
```

```
## [1] 15.456148  3.791572
```

```r
# Erro
fit$erro
```

```
## [1] 5.580005e-07
```

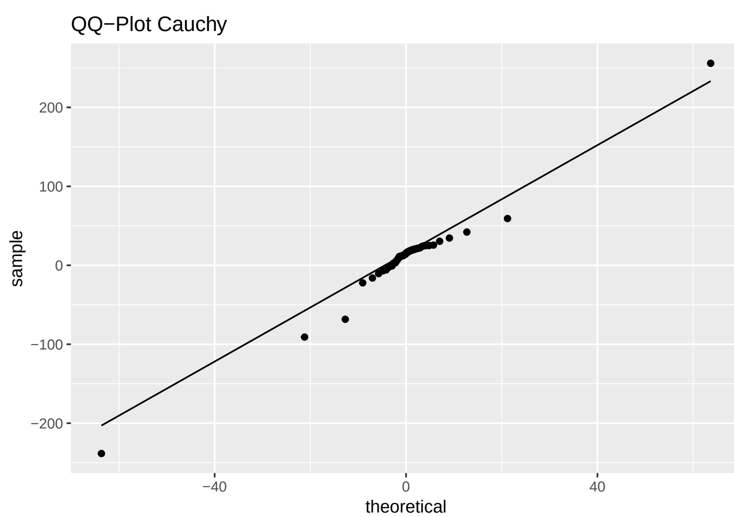
```r
# Verificando se os autovalores são negativos
eigen(fit$H) # Autovalores são estritamente negativos -> ponto de máximo
```

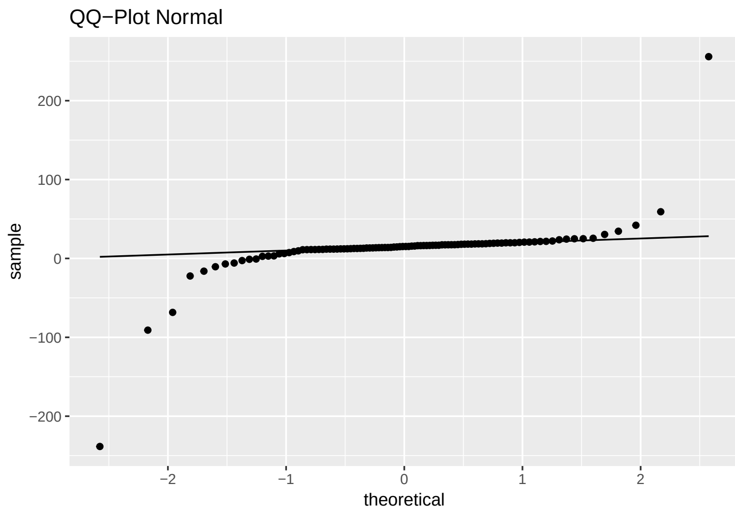
```
## eigen() decomposition
## $values
## [1] -2.949971 -8.432503
##
## $vectors
##             [,1]        [,2]
## [1,] -0.99990942 -0.01345892
## [2,] -0.01345892  0.99990942
```

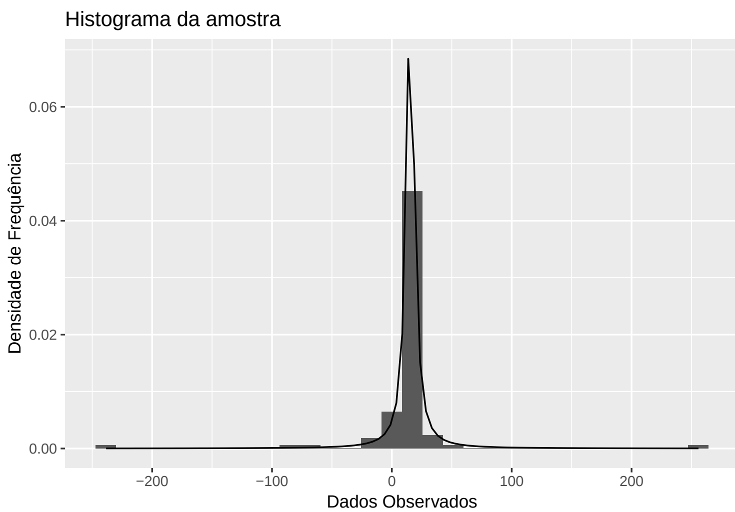
```r
# Estimativas de máxima verossimilhança
t1.hat <- fit$theta[1]
t2.hat <- fit$theta[2]
```

(c) Justifique o uso de uma distribuição Cauchy para modelagem estatística do conjunto de dados abaixo. Depois, encontre as estimativas de máxima verossimilhança dos parâmetros de localização μ e escala σ^2.

Resolução







Como temos dados em um range de -238.41 à 255.83, notei que provem de alguma distribuição que tem seu suporte nos $\mathbb{R}$ e como obtive estimativas positivas pensei na distribuição Normal ou Cauchy, através dos gráficos QQ e do histograma com a densidade da distribuição de Cauchy com os parâmetros obtidos através do item 2-b, as estimativas se ajustam muito bem ao histograma da amostra. Assim, com o algoritmo do exercício anterior, temos que:

$$\begin{pmatrix} \hat{\mu}_{mv} \\ \hat{\sigma}_{mv} \end{pmatrix} = \begin{pmatrix} 15.456 \\ 3.791 \end{pmatrix}$$

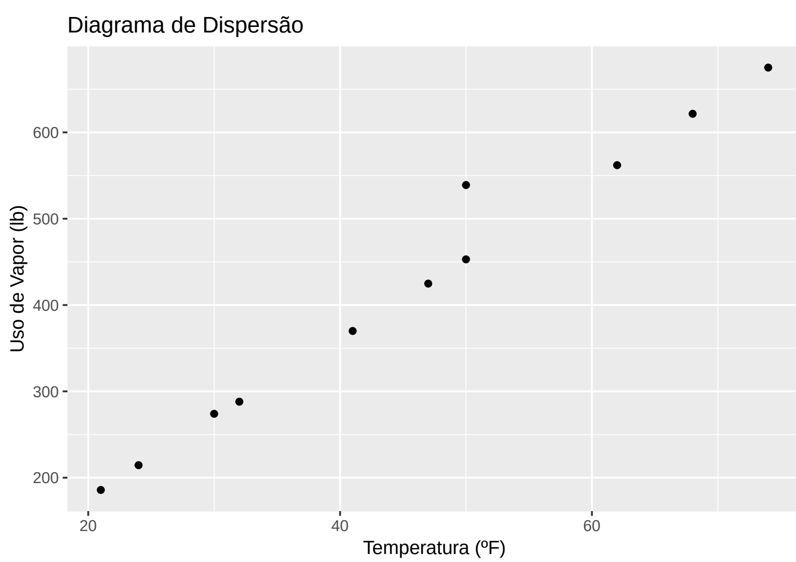
Questão 3

Considere o conjunto de dados abaixo referente ao uso de vapor (em libras) por uma indústria em função da temperatura média mensal (em graus Fahreinheit).

Temperatura (ºF)	Uso de vapor (lb)
21	185.79
24	214.47
32	288.03
47	424.84
50	539.03
68	621.55
74	675.06
62	562.03
50	452.93
41	369.95
30	273.98

(a) Faça um gráfico de dispersão e justifique o uso de um modelo de regressão linear simples. Formalize tal modelo e suas suposições.

Resolução



O modelo de regressão linear simples se adequa por termos cada observação independente das outras, além disso obervando o gráfico de disperção podemos observar uma tendência linear, para verificar tal observação foi calculado o coeficiente de correlação linear de Pearson entre as variáveis uso de vapor e temperatura que resultou em 0.988, e como está muito próximo de 1 é possível afirmar que há uma forte correlação linear positiva entre as variáveis. Além do mais, para a utilização de tal modelo $Y_i = \beta_0 + \beta_1 x + e_i$ supus que $\mathbb{E}(e_i) = 0$ e $Var(e_i) = \sigma^2$ e também homocedasticidade entre os erros.

(b) Estime β_0 e β_1 pelo método de mínimos quadrados e interprete as estimativas obtidas

Resolução

Como visto em aula os estimadores de mínimos quadrados para β_0 e β_1 são:

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}}$$

Onde S_{xy} é a covariância amostral de $(x_1, ..., x_n)$ e $(y_1, ..., y_n)$ e S_{xx} é a variância amostral de $(x_1, ..., x_n)$, assim as estimativas de mínimos quadrados são:

$$\widehat{\beta_0}_{mq} = 418.88 - 9.324 * 45.3636 = -4.09$$

$$\widehat{\beta_1}_{mq} = \frac{2926.415}{313.8545} = 9.324$$

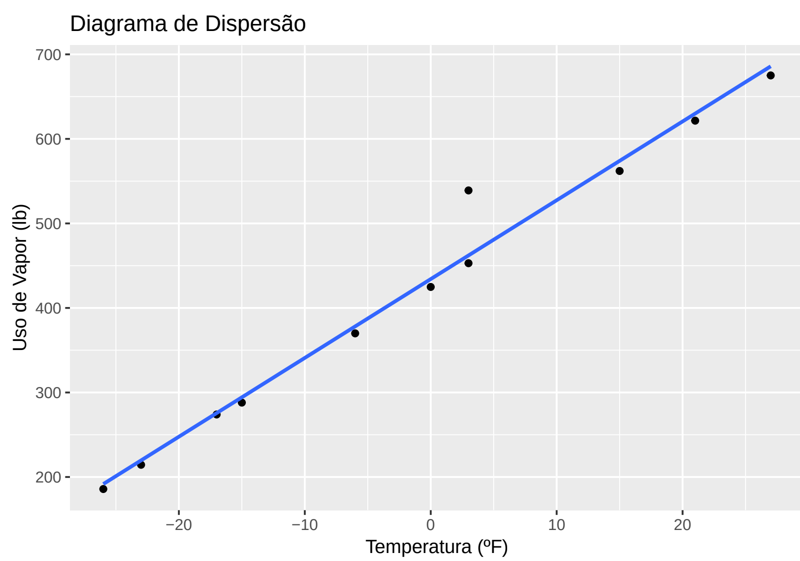
Logo temos que a equação do modelo é $Y_i = -4.09 + 9.324x_i$ onde $\widehat{\beta_1}_{mq} = 9.324$ pode ser interpretado como o acréscimo esperado de 9.324 lb no uso do vapor quando a temperatura aumenta em 1 ºF na média mensal. E $\widehat{\beta_0}_{mq} = -4.09$ pode ser interpretado como o valor esperado do uso de vapor onde a temperatura média mensal seja 0 ºF, o uso de vapor deveria ser -4.097, entretanto esse valor não faz sentido por ser negativo, para isso foi proposto um deslocamento dos dados de temperatura subtraindo cada valor de sua mediana amostral, assim reajustando o modelo obtemos:

$$Y_i = 434.136 + 9.324 * (x - 47)$$

onde foram mantidas todas as suposições anteriores, assim é possível interpretar $\widehat{\beta_0}_{mq} = 434.136$ como sendo o valor esperado de 434.136 lb quando a temperatura é 0 ºF.

(c) Adicione ao gráfico do item (a) a curva de regressão estimada. Qual é a quantidade esperada de vapor a ser utilizada pela indústria em um mês com temperatura média de 70 ºF?

Resolução



Usando a equação para prever a quantidade esperada de vapor a ser utilizada pela indústria em um mês com temperatura média de 70 ºF é: $Y = 434.136 + 9.324 * (70 - 47) \Rightarrow Y = 648.59$. Ou seja, o valor é aproximadamente: 648.59 lb.