Universidade Federal de Minas Gerais

Homework 1

A.P. Braga

August 26, 2019

Aproximação Polinomial

0.1 Introdução Teórica

Considere um polinômio de grau p conforme representado na sua forma geral na Equação 0.1.

$$p(x) = w_p x^p + w_{p-1} x^{p-1} + \cdots + w_1 x + w_0 \tag{0.1}$$

em que x é o argumento e w_i é o coeficiente do termo de grau i.

Dadas as observações (x_i, y_i) representadas na forma do conjunto de dados $D = \{x_i, y_i\}_{i=1}^{N}$, deseja-se encontrar o polinômio de grau p que melhor aproxima a função geradora $f_g(x)$ do conjunto D. O objetivo é, a partir das amostras de dados, encontrar o grau p e os coeficientes w_i de forma tal que $p(x) \approx f_g(x)\ \forall x$. A aproximação de $f_g(x)$ é usualmente feita com base na minimização do erro dos termos quadráticos $(y_i - p(x_i))^2$ $(i = 1 \cdots N)$. Espera-se que o conjunto D contenha informação suficiente para que seja possével aproximar $f_g(x)$ por $p(x)$ com base somente nas suas N amostras. Os parâmetros de $p(x)$ são ajustados de forma tal que $y_i = w_p x_i^p + w_{p-1} x_i^{p-1} + \cdots + w_1 x_i + w_0\ \forall x_i \in D$, conforme representado no sistema de equações 0.2.

$$\begin{array}{lllllll}
y_1 = w_p x_1^p & +w_{p-1} x_1^{p-1} & + & \cdots & +w_1 x_1 & +w_0 \\
y_2 = w_p x_2^p & +w_{p-1} x_2^{p-1} & + & \cdots & +w_1 x_2 & +w_0 \\
\vdots & \vdots & \cdots & \vdots & & \\
y_N = w_p x_N^p & +w_{p-1} x_N^{p-1} & + & \cdots & +w_1 x_N & +w_0
\end{array} \tag{0.2}$$

O sistema representado em 0.2 possui N equações e p incógnitas, podendo também ser representado na forma matricial 0.3.

$$\mathbf{Hw} = \mathbf{y} \tag{0.3}$$

em que $\mathbf{H}$, $\mathbf{w}$ e $\mathbf{y}$ são representados em 0.4, 0.5 e 0.6.

$$\mathbf{H} = \begin{bmatrix} x_1^p & x_1^{p-1} & \cdots & x_1 & 1 \\ x_2^p & x_2^{p-1} & \cdots & x_2 & 1 \\ \vdots & \vdots & \cdots & \vdots & \\ x_N^p & x_N^{p-1} & \cdots & x_N & 1 \end{bmatrix} \tag{0.4}$$

$$\mathbf{w} = \begin{bmatrix} w_p \\ w_{p-1} \\ \vdots \\ w_1 \\ w_0 \end{bmatrix} \tag{0.5}$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \tag{0.6}$$

A matriz $\mathbf{H}$ possui um papel importante na resolução do problema de aproximação, pois ela contém os termos não-lineares que compõem o polinômio $p(x)$, os quais serão responsáveis pela projeção dos elementos x_i no espaço composto pelo sistema de coordenadas caracterizado pelas colunas de $\mathbf{H}$. Como $\mathbf{H}$ e $\mathbf{y}$ são dados pelo problema, a solução da Equação 0.3 pode ser obtida por meio da pseudoinversa, conforme Equação 0.7.

$$\mathbf{w} = \mathbf{H}^{+}\mathbf{y} \tag{0.7}$$

em que $\mathbf{H}^{+}$ é a pseudoinversa de $\mathbf{H}$.

0.1.1 Aproximação de uma Função Quadrática

```
> rm(list=ls())
> library('corpcor')
> # Geração dos dados
> N<-20
> x<-runif(n = N,min=-15,max=10)
> xgrid<-seq(-15,10,0.1)
> yr<-(0.5*x^2+3*x+10) + 10*rnorm(length(x))
> ygrid<-(0.5*xgrid^2+3*xgrid+10)
> # Aproximação de grau dois
> H<-cbind(x^2,x,1)
```

```
> w<-pseudoinverse(H) %*% yr
> # Aproximação
> Hgrid<-cbind(xgrid^2,xgrid,1)
> yhat<-H %*% w
> yhatgrid<-Hgrid %*% w
>
> # plotando a projeção
>
```

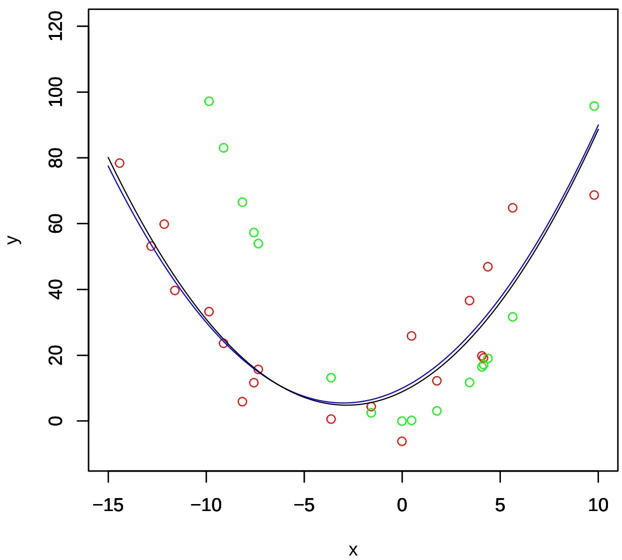

Figure 0.1: Aproximação de uma função por meio de um polinômio de grau 2. Os dados foram gerados a partir da função $f_g(x) = \frac{1}{2}x^2 + 3x + 10$, os quais são mostrados na cor vermelha na figura. A linha contínua mostra a função aproximada. Os dados em preto representam as projeções de x no plano cartesiano formado pelas duas primeiras colunas da matriz $\mathbf{H}$.

0.1.2 Sub-ajuste (Underfitting)

```
> rm(list=ls())
> library('corpcor')
> # Geração dos dados
> N<-20
> x<-runif(n = N,min=-15,max=10)
```

```
> xgrid<-seq(-15,10,0.1)
> yr<-(0.5*x^2+3*x+10) + 10*rnorm(length(x))
> ygrid<-(0.5*xgrid^2+3*xgrid+10)
> # Aproximação de grau um
> H<-cbind(x,1)
> w<-pseudoinverse(H) %*% yr
> # Aproximação
> Hgrid<-cbind(xgrid,1)
> yhatgrid<-Hgrid %*% w
>
```

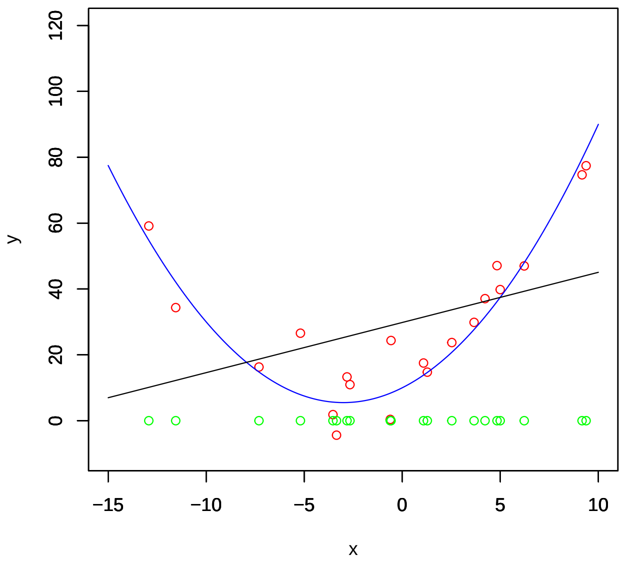

Figure 0.2: Aproximação com função de grau inferior resulta em *underfitting*.

0.1.3 Sobre-ajuste (Overfitting)

```
> rm(list=ls())
> library('corpcor')
> # Geração dos dados
> N<-20
> x<-runif(n = N,min=-15,max=10)
> xgrid<-seq(-15,10,0.1)
> yr<-(0.5*x^2+3*x+10) + 10*rnorm(length(x))
```

```
> ygrid<-(0.5*xgrid^2+3*xgrid+10)
> # Aproximação de grau maior do que 2
> H<-cbind(x^10,x^9,x^8,x^7,x^6,x^5,x^4,x^3,x^2,x,1)
> w<-pseudoinverse(H) %*% yr
> # Aproximação
> Hgrid<-cbind(xgrid^10,xgrid^9,xgrid^8,xgrid^7,
   xgrid^6,xgrid^5,xgrid^4,xgrid^3,xgrid^2,xgrid,1)
> yhatgrid<-Hgrid %*% w
```

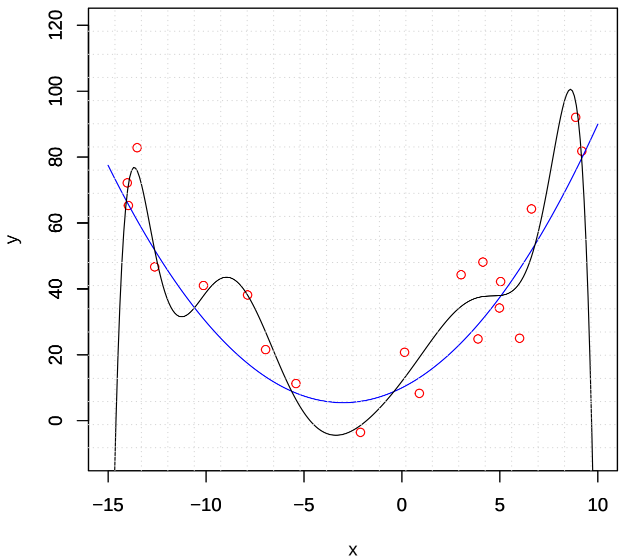

Figure 0.3: Aproximação com função de grau superior resulta em *overfitting*.

0.1.4 Efeito do Tamanho do Conjunto de Amostras

```
> rm(list=ls())
> library('corpcor')
> # Geração dos dados
> N<-1000
> x<-runif(n = N,min=-15,max=10)
> xgrid<-seq(-15,10,0.1)
> yr<-(0.5*x^2+3*x+10) + 10*rnorm(length(x))
> ygrid<-(0.5*xgrid^2+3*xgrid+10)
> # Aproximação de grau maior do que 2
```

```
> H<-cbind(x^5,x^4,x^3,x^2,x,1)
> w<-pseudoinverse(H) %*% yr
> # Aproximação
> Hgrid<-cbind(xgrid^5,xgrid^4,xgrid^3,xgrid^2,xgrid,1)
> yhatgrid<-Hgrid %*% w
```

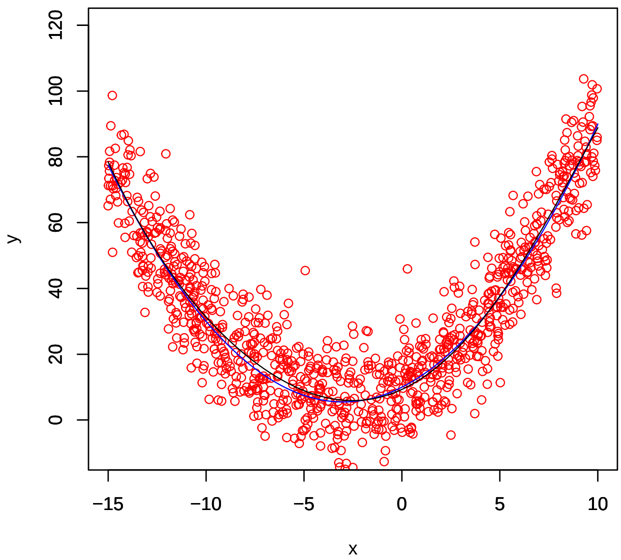

Figure 0.4: Efeito do tamanho da amostra na aproximação quando o modelo é superdimensionado.

1 Exercícios

Considerando a aproximação polinomial das seções anteriores, faça:

1. Para um número N de amostras, por exemplo $N = 20$, plote o gráfico de erro médio quadrático em relação à função geradora $f_g(x) = \frac{1}{2}x^2 + 3x + 10$ quando a grau do polinômio varia de $p = 1$ a $p = 10$.

2. Para a mesma faixa de valores de graus de polinômios acima, apresente o gráfico da variância do erro das amostras em relação a $f_g(x) = \frac{1}{2}x^2 + 3x + 10$ em função do grau do polinômio.

3. Repita o processo do item 1 vinte vezes, capturando o erro para os 10 valores diferentes de p a cada iteração. Calcule então a variância do erro em função do grau do polinômio.

Para o cálculo da pseudo-inversa o aluno deverá usar o pacote *library("corpcor")*. E para a geração do gráfico do tipo *boxplot* deverá ser usada a biblioteca *library ('boxplot-dbl')*. A seguir segue o modelo de resposta para o exercício 1 e 2 (a resposta não será necessariamente igual).

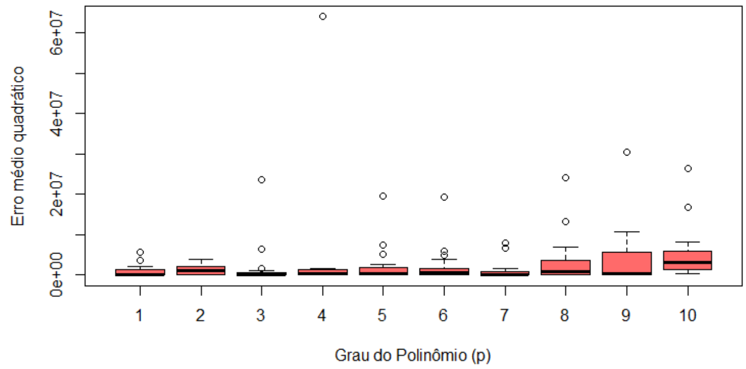

Figure 1.1: Erro médio quadrático em relação à função geradora x Grau do Polinômio

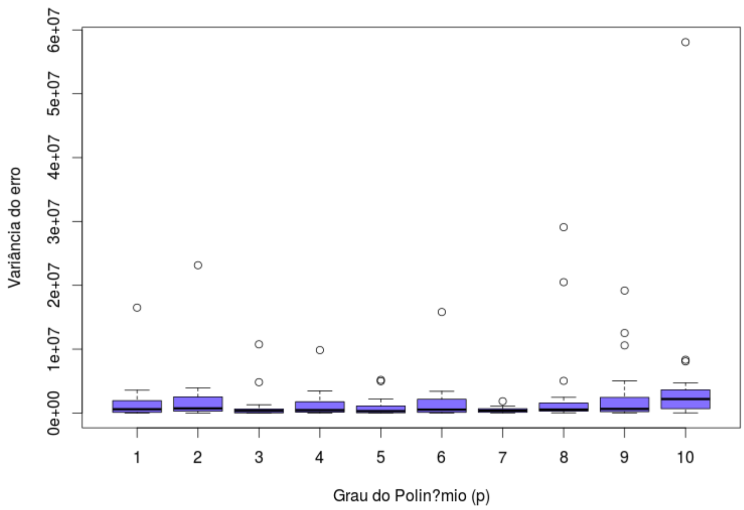

Figure 1.2: Variância do erro em relação à função geradora x Grau do Polinômio