

# Estudo de crescimento da população do País 14 com dados referentes aos anos de 1960 a 2013.<sup>1</sup>

Henrique Pasquini Santos<sup>2</sup> – nº USP 8532252

Raul Zaninetti Rosa<sup>3</sup> – nº USP 8517310

Universidade de São Paulo, São Carlos, SP

## INTRODUÇÃO

Neste artigo, teremos como objetivo determinar, com base em pontos fornecidos sobre a população do País 14 entre os anos de 1960 a 2013, uma função  $f(x)$  que melhor se aproxima da função definida pelos pares de pontos. Para encontrar a função  $f(x)$ , será utilizado o Método dos Mínimos Quadrados com diferentes graus. A partir dos resultados obtidos no método, será verificado qual grau teve a melhor aproximação, e a partir dessa aproximação será estimado a população do País 14 para o ano de 2018. Por fim, será discutido a qualidade da função de aproximação obtida com relação aos dados fornecidos e a qualidade da aproximação para a população estimada em 2018.

## PROCEDIMENTOS

Existem diferentes formas de calcular uma função  $f(x)$  através de pares de pontos, que são: a Interpolação Polinomial e o Método dos Mínimos Quadrados. No entanto, a Interpolação Polinomial é um o método que gera um polinômio de grau  $N-1$ , sendo  $N$  o número de pontos utilizados. Sabendo que será utilizada uma grande quantidade de pontos, vemos que a quantidade máxima de raízes é  $N-1$ , tornando o método muito custoso e inconveniente para a resolução deste problema, portanto ele deve ser descartado. Logo, resta apenas o Método dos Mínimos Quadrados, o qual pode gerar um polinômio de qualquer grau, o que é útil neste caso, pois serão calculados e analisados polinômios com diferentes graus e será discutido qual deles possui uma melhor aproximação da função definida pelos pares de pontos. Como são dados pontos, será utilizado o caso discreto. Além disso, serão calculados e discutidos os erros de truncamento da função que são utilizados como referência para saber qual aproximação é a mais próxima da função ideal.

---

<sup>1</sup> Trabalho realizado durante a disciplina SME0104 – ICMC – USP São Carlos, SP, 2 de junho de 2015

<sup>2</sup> Estudante de Graduação, 5º Semestre do Curso de Bacharelado em Ciências da Computação, USP São Carlos, SP; e-mail: [riquepasq@gmail.com](mailto:riquepasq@gmail.com)

<sup>3</sup> Estudante de Graduação, 5º Semestre do Curso de Bacharelado em Ciências da Computação, USP São Carlos, SP; e-mail: [raulzrosa@gmail.com](mailto:raulzrosa@gmail.com)

O método dos Mínimos Quadrados é definido como:

Uma função é dada por uma tabela com  $n+1$  pares de pontos  $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ , onde  $y_i = f(x_i)$ ,  $i = 0, 1, \dots, n$ , com os  $n+1$  pontos  $x_0, x_1, \dots, x_n$  distintos.

A partir desses pontos vamos determinar o polinômio

$$P_m(x) = a_0 + a_1 x + \dots + a_m x^m$$

de grau no máximo  $m$ , com  $m \leq n$ , tal que o erro de truncamento  $Q$  seja mínimo.

A variável  $Q$  é definida por

$Q = \sum_{k=0}^n |f(x_k) - P_m(x_k)|^2$ , usando o produto escalar obtemos:

$$Q = \sum_{k=0}^n (y_k - P_m(x_k))^2.$$

O método consiste então em usar os vetores

$$\vec{y} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}; \quad \vec{p} = \begin{pmatrix} P_m(x_0) \\ P_m(x_1) \\ \vdots \\ P_m(x_n) \end{pmatrix} = a_0 \vec{u}_0 + a_1 \vec{u}_1 + \dots + a_m \vec{u}_m,$$

com

$$\vec{u}_0 = (1 \quad \dots \quad 1)^T; \quad \vec{u}_k = (x_0^k \quad \dots \quad x_n^k)^T, \quad k = 1, 2, \dots, m,$$

resolver o sistema linear abaixo a fim de achar  $a_0, a_1, \dots, a_m$ .

$$\begin{pmatrix} (\vec{u}_0, \vec{u}_0) & (\vec{u}_1, \vec{u}_0) & \dots & (\vec{u}_m, \vec{u}_0) \\ (\vec{u}_0, \vec{u}_1) & (\vec{u}_1, \vec{u}_1) & \dots & (\vec{u}_m, \vec{u}_1) \\ \vdots & \vdots & \ddots & \vdots \\ (\vec{u}_0, \vec{u}_m) & (\vec{u}_1, \vec{u}_m) & \dots & (\vec{u}_m, \vec{u}_m) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} (\vec{y}, \vec{u}_0) \\ (\vec{y}, \vec{u}_1) \\ \vdots \\ (\vec{y}, \vec{u}_m) \end{pmatrix}$$

Considerando então a função  $y = f(x)$  dada pela Tabela 1, temos que  $y$  é a população referente ao ano  $x$ .

Tabela 1 - População do País 14 durante os anos 1960-2013

x	y	x	y
1960	498823	1987	872158
1961	503762	1988	896429
1962	509348	1989	921314
1963	515762	1990	946703
1964	523236	1991	972539
1965	531901	1992	998823
1966	541895	1993	1025559
1967	553098	1994	1052772
1968	565080	1995	1080477
1969	577241	1996	1108698
1970	589165	1997	1137412
1971	600692	1998	1166525
1972	611990	1999	1195919
1973	623375	2000	1225527
1974	635318	2001	1255299
1975	648174	2002	1285318
1976	662036	2003	1315820
1977	676829	2004	1347125
1978	692535	2005	1379465
1979	709092	2006	1412907
1980	726454	2007	1447388
1981	744624	2008	1482843
1982	763639	2009	1519155
1983	783527	2010	1556222
1984	804319	2011	1594034
1985	826025	2012	1632572
1986	848654	2013	1671711

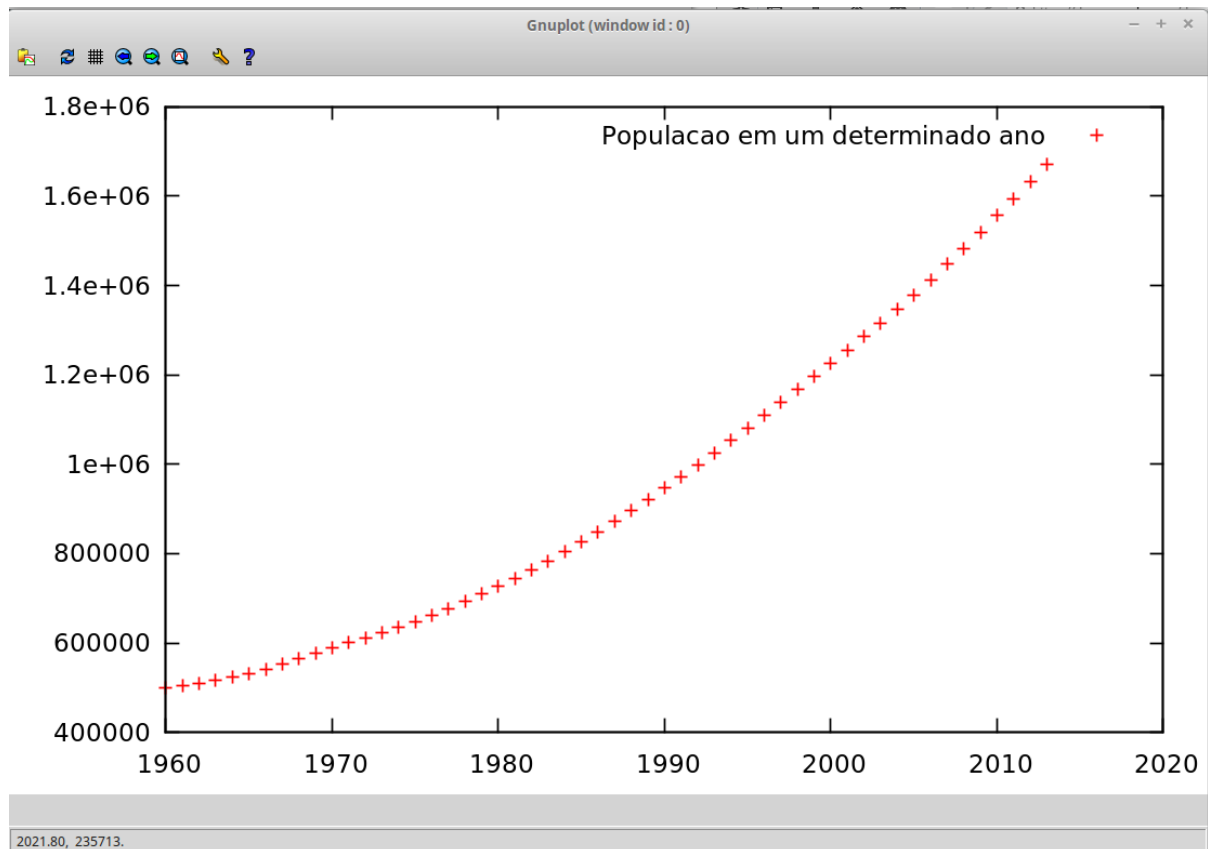
Através destes dados, podemos gerar um gráfico, utilizando a ferramenta GNUPLOT (ferramenta *open souce* do Linux):

Comandos:

*\$ gnuplot*

*\$ plot "pontos.dat" using 1:2 title 'Populacao em um determinado ano'*

Gráfico 1 - Função definida pelos pares de pontos referentes à Tabela 1



Foi implementado um programa na linguagem C que, dado o número de pontos e o grau do polinômio desejado, são feitos os cálculos dos coeficientes  $(u_i, u_j)$  e  $(y, u_i)$ , porém não resolve o sistema linear resultante. O programa também calcula o erro de truncamento para cada um dos graus que serão analisados e a estimativa da população para o ano de 2018.

## RESULTADOS

Iremos aplicar então o Método dos Mínimos Quadrados para vários graus diferentes a fim de encontrar uma função que mais de aproxima do gráfico plotado acima.

Vamos aplicar o método primeiramente para polinômio do tipo  $P_1(x) = a_0 + a_1 \cdot x$ .

Definindo os vetores

$y = (498823, 503762, 509348, \dots, 1671711);$

$u_0 = (1, 1, \dots, 1);$

$u_1 = (1960, 1961, \dots, 2013),$

Obtemos o sistema linear

$$\begin{pmatrix} (\vec{u}_0, \vec{u}_0) & (\vec{u}_1, \vec{u}_0) \\ (\vec{u}_0, \vec{u}_1) & (\vec{u}_1, \vec{u}_1) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} (\vec{y}, \vec{u}_0) \\ (\vec{y}, \vec{u}_1) \end{pmatrix}.$$

Utilizando-se do programa em C, colocamos como entrada N = 54 (número de pontos) e o grau polinomial igual à 1.

$$\begin{pmatrix} 54 & 107271 \\ 107271 & 213106959 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 50767308 \\ 101140130242 \end{pmatrix}$$

Calculou-se o sistema linear através do site “<http://www.wolframalpha.com>”, obtendo a seguinte função linear:

$$y = -\frac{226194716246}{5247} + \frac{116349160}{5247} x$$

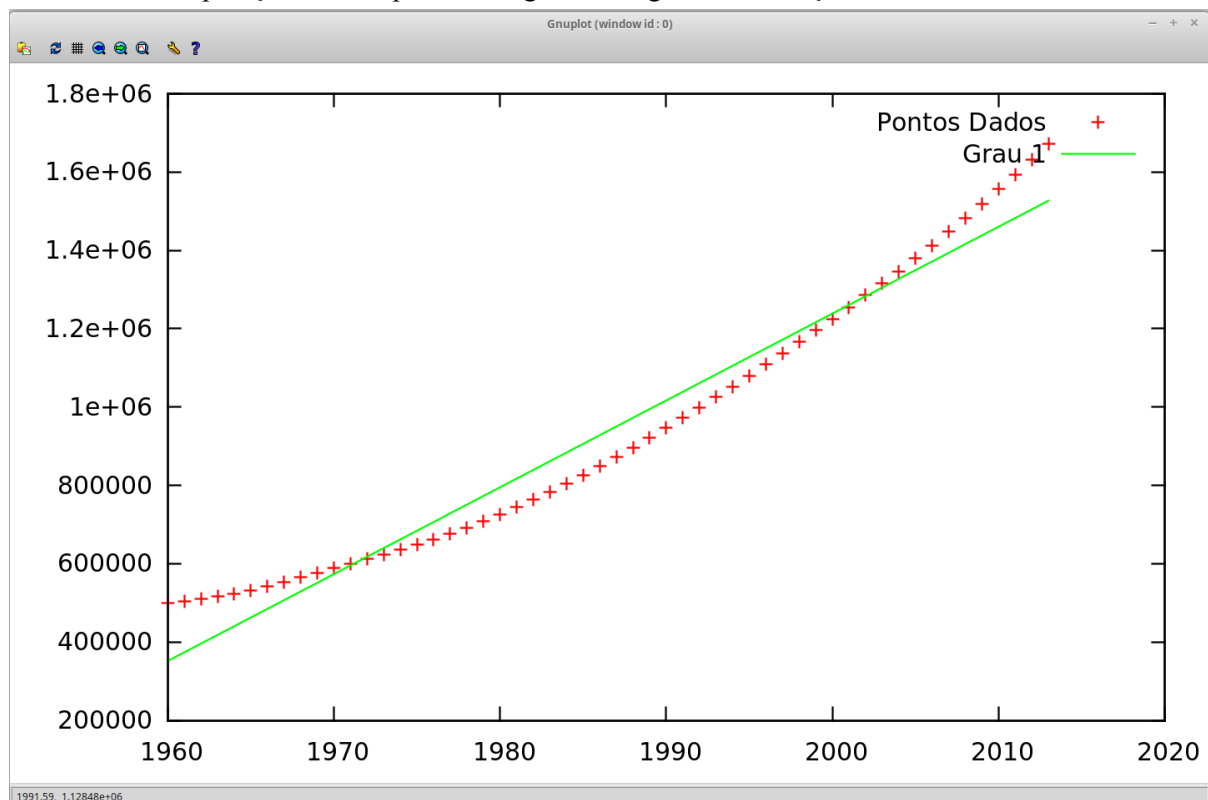
Plotando a função no GNUPLOT obtemos o seguinte gráfico:

Comandos:

`$ gnuplot`

`$ plot "teste" using 1:2 title 'Pontos Dados', -226194716246/5247 + 116349160/5247*x title 'Grau 1'`

Gráfico 2 - Comparação entre o polinômio gerado de grau 1 e a função do Gráfico 1



Através do programa em C, iremos calcular também o erro de truncamento Q:

$$Q = ||f - P_1||^2 = \sum_{k=0}^{54} (y_k - P_1(x_k))^2$$

$$Q = 254712884661.501068$$

Agora, será aplicado o método para um polinômio do tipo  $P_2(x) = a_0 + a_1x + a_2x^2$ .

Definindo os vetores:

$$y = (498823, 503762, 509348, \dots, 1671711);$$

$$u_0 = (1, 1, \dots, 1);$$

$$u_1 = (1960, 1961, \dots, 2013),$$

$$u_2 = (1960^2, 1961^2, \dots, 2013^2)$$

$$\begin{pmatrix} (u_0, u_0) & (u_1, u_0) & (u_2, u_0) \\ (u_0, u_1) & (u_1, u_1) & (u_2, u_1) \\ (u_0, u_2) & (u_1, u_2) & (u_2, u_2) \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} (y, u_0) \\ (y, u_1) \\ (y, u_2) \end{pmatrix}.$$

Utilizando-se do programa em C, colocamos como entrada N = 54 (número de pontos) e o grau polinomial igual à 2.

$$\begin{pmatrix} 54 & 107271 & 213106959 \\ 107271 & 213106959 & 423389089881 \\ 213106959 & 423389089881 & 841217724918603 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 50767308 \\ 101140130242 \\ 201505824035820 \end{pmatrix}$$

Calculou-se o sistema linear através do site "<http://www.wolframalpha.com>", obtendo a seguinte função de grau 2:

$$y = \frac{765756916521955}{636636} - \frac{4707147028939}{3819816}x + \frac{36548591}{115752}x^2$$

Plotando a função no GNUPLOT obtemos o seguinte gráfico:

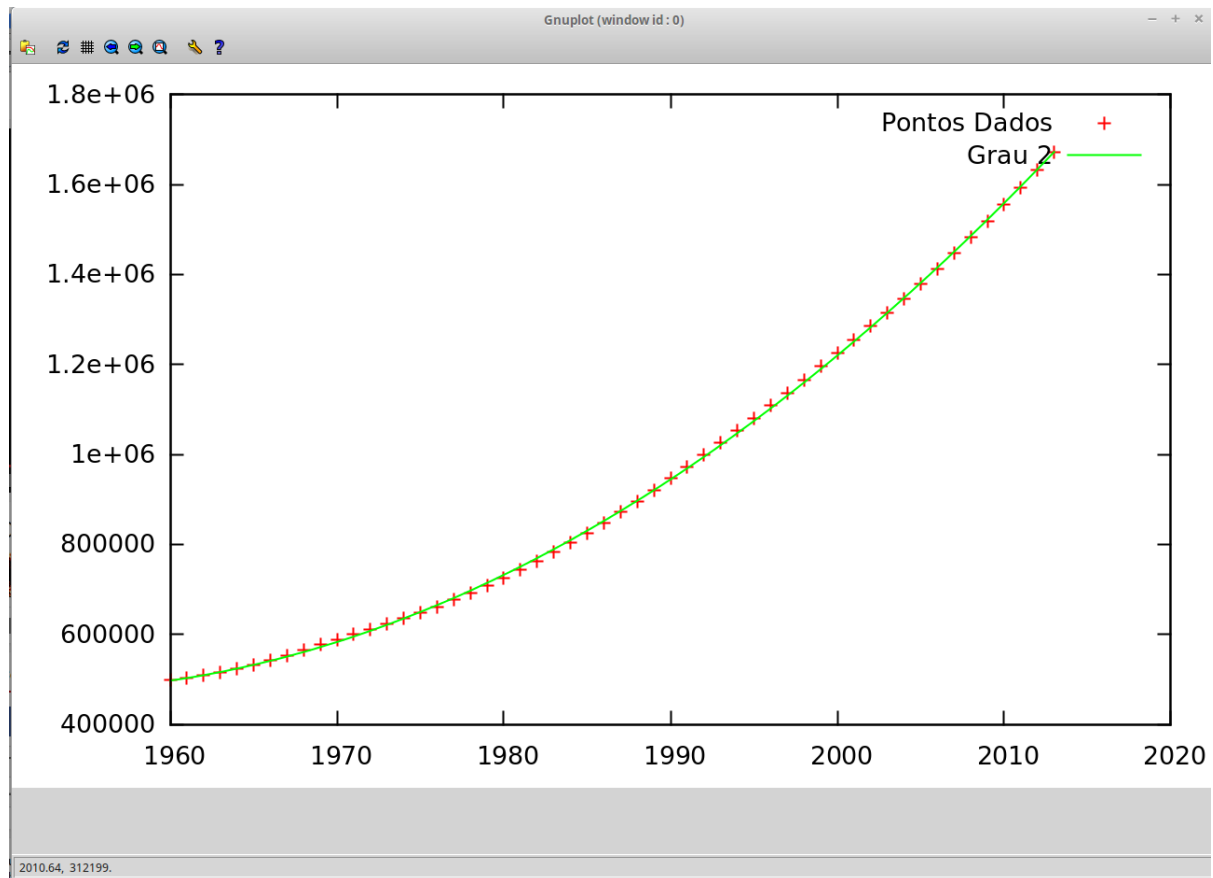
Comandos:

`$ gnuplot`

`$ plot "teste" using 1:2 title 'Pontos Dados', 765756916521955/636636 -`

`x*4707147028939/3819816 + x*x*36548591/115752 title 'Grau 2'`

Gráfico 3 - Comparação entre o polinômio gerado de grau 2 e a função do Gráfico 1



Iremos calcular também o erro de truncamento Q:

$$Q = ||f - P_1||^2 = \sum_{k=0}^{54} (y_k - P_1(x_k))^2$$

$$Q = 828812767.487610$$

É perceptível que o erro de truncamento para esse caso é significativamente menor que o do caso anterior. Além disso, vemos que, ao analisar o gráfico, a função encontrada nesse caso é muito semelhante à função definida pelos pares de pontos.

Após isso, será calculado o polinômio  $P_3(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ . Através do programa em C, temos:

$$\begin{pmatrix} 54 & 107271 & 213106959 & -40467378087 \\ 107271 & 213106959 & 423389089881 & -80233148699829 \\ 213106959 & 423389089881 & 841217724918603 & -159084465697926940 \\ -40467378087 & -80233148699829 & -159084465697926940 & 32164746247387099000 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 50767308 \\ 101140130242 \\ 201505824035820 \\ -34596380864273780 \end{pmatrix}$$

Calculou-se o sistema linear através do site “<http://www.wolframalpha.com>”, obtendo a seguinte função de grau 3:

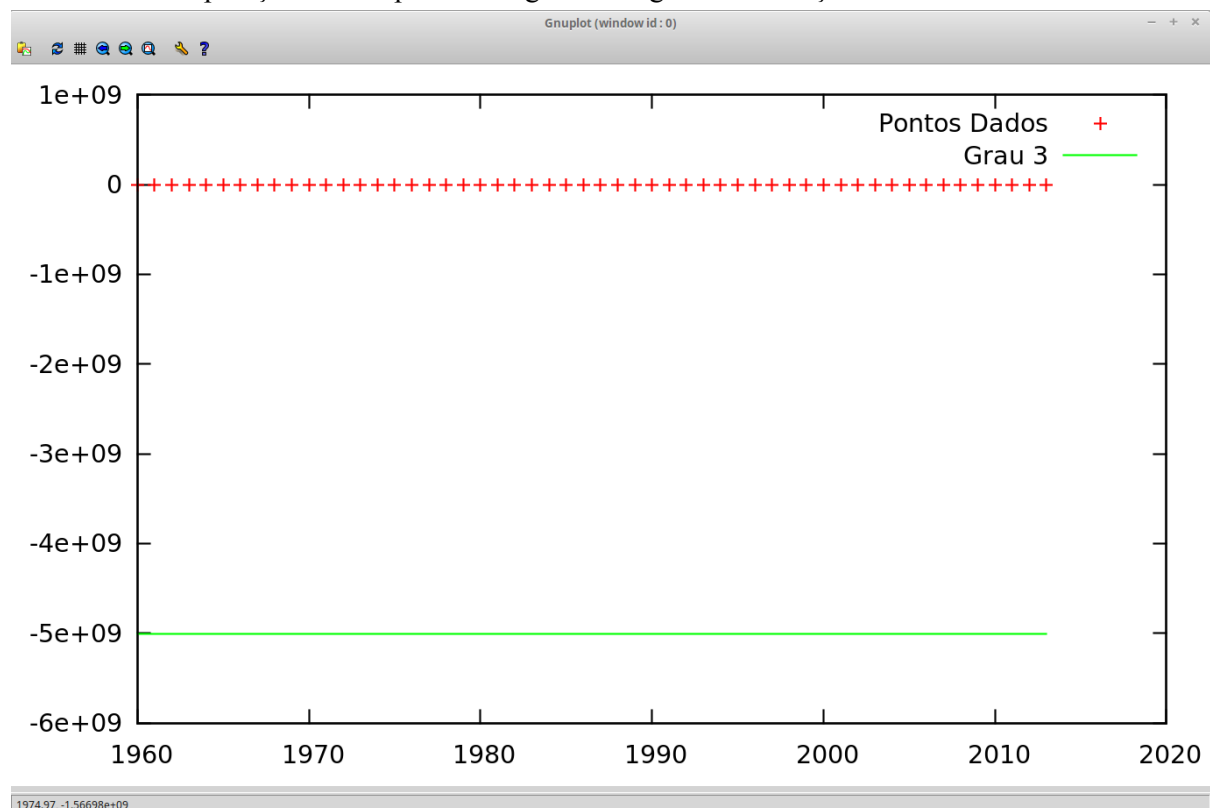
$$y = \frac{24318339644497207467959330962}{31804099838240572557} - \frac{258662008879128841715623106}{31804099838240572557} x + \frac{22970630670017999848}{6061387428671731} x^2 - \frac{271792974385790}{466260571436287} x^3$$

Plotando a função no GNUPLOT obtemos o seguinte gráfico:

Comandos:

```
$ gnuplot
$plot "teste" using 1:2 title 'Pontos Dados',
24318339644497207467959330962/31804099838240572557 -
x*258662008879128841715623106/31804099838240572557 +
x*x*22970630670017999848/6061387428671731 -
x*x*x*271792974385790/466260571436287 title 'Grau 3'
```

Gráfico 4 - Comparação entre o polinômio gerado de grau 3 e a função do Gráfico 1



Iremos calcular também o erro de truncamento Q:

$$Q = ||f - P_1||^2 = \sum_{k=0}^{54} (y_k - P_1(x_k))^2$$

$$Q = 1353918566499902029824.000000$$



Após o cálculo do erro de truncamento Q para a função de grau 3, foi visto que o erro gerado apresenta um valor muito maior que no caso anterior. Logo, analisando estes resultados, vimos que não é necessário aplicar o método para polinômios de graus maiores, pois à partir do grau 3 em diante este erro aumenta cada vez mais e o método já não é mais otimizado.

Portanto, a função escolhida que mais se aproxima dos dados é a quadrática. À partir disso, será estipulada a população do país para o ano de 2018.

Utilizando o polinômio quadrático:

$$y = \frac{765756916521955}{636636} - \frac{4707147028939}{3819816} x + \frac{36548591}{115752} x^2$$

Com  $x = 2018$ , temos:

$$y = \frac{765756916521955}{636636} - \frac{4707147028939}{3819816} 2018 + \frac{36548591}{115752} 2018^2$$
$$y = \mathbf{1875230,767816}$$

A população estimada para o ano 2018 é de aproximadamente 1875230 habitantes.

## CONCLUSÕES

Vimos que a função que melhor se aproxima dos dados é a quadrática, pois foi a que obteve o menor erro de truncamento e visualmente também é a que mais se aproxima da curva da função definida pelos pontos.

Em relação ao polinômio de grau 1, apesar de não apresentar um erro muito grande e possuir alguns pontos semelhantes aos dos dados (Gráfico 2), ele é impreciso e, quando comparado ao de grau 2, passa a apresentar uma qualidade muito inferior à deste.

Já o polinômio de grau 3 apresenta valores inviáveis para se manipular, com um valor de erro extremamente alto, e sua curva muito distante da curva dos dados (Gráfico 4).

Logo, podemos concluir que o resultado atingido para a população estimada para o ano de 2018 (à partir da função quadrática) é plausível, visto que pode-se inferir, pelo Gráfico 3, que as curvas de ambas as funções estão tendendo para aproximadamente 1875230 habitantes em 2018.

## REFERÊNCIAS

RUGGIERO, M. A. G.; LOPES, V. L. R. L. Cálculo Numérico Aspectos Teóricos e Computacionais. 2. ed. , Capítulo 6. São Paulo: Makron, 1997.  
FRANCO, N. B. Cálculo Numérico, Capítulo 7. São Paulo: Pearson Prentice Hall, 2006.