Modelagem da Linha 9 - Esmeralda da CPTM utilizando o Método dos Quadrados Mínimos e a Série de Fourier

Karl Willian Sousa Santos

Trabalho de Conclusão do Curso Superior de Licenciatura em Matemática, orientado pelo Prof. Me. Lucas Casanova Silva

IFSP
São Paulo

2016

Dados Internacionais de Catalogação na Publicação (CIP)

Santos, Karl Willian Sousa

Modelagem da Linha 9 - Esmeralda utilizando o Método dos Quadrados Mínimos e a Série de Fourier / Karl Willian Sousa Santos - São Paulo: IFSP, 2016.

79 f.

Trabalho de Conclusão do Curso Superior de Licenciatura em Matemática - Instituto Federal de Educação, Ciência e Tecnologia de São Paulo

Orientador: Lucas Casanova Silva

1. Modelagem. 2. CPTM. 3. Método dos Quadrados Mínimos. 4. Série de Fourier. I. Modelagem da Linha 9 - Esmeralda utilizando o Método dos Quadrados Mínimos e a Série de Fourier.

KARL WILLIAN SOUSA SANTOS

MODELAGEM DA LINHA 9 – ESMERALDA DA CPTM UTILIZANDO O MÉTODO DOS QUADRADOS MÍNIMOS E A SÉRIE DE FOURIER

Monografia apresentada ao Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, em cumprimento ao requisito exigido para a obtenção do grau acadêmico de Licenciado em Matemática.

APROVADA EM 1º/07/2016 **CONCEITO:** 9,5

Mariana P. M. A. Baroni

Profa. Dra. Mariana Pelissari Monteiro Aguiar Baroni

Membro da Banca

Prof. Dr. Marco Aurélio Granero Santos

Membro da Banca

Lucas Casanova Silva

Prof. Me. Lucas Casanova Silva

Orientador

Karl Willian Sousa Santos

Aluno: Karl Willian Sousa Santos

Agradecimentos

Agradeço aos meus pais, Regiane e Carlos, grandes incentivadores, pela determinação e luta na minha formação, pela paciência, apoio e confiança.

Aos meus familiares pela compreensão e apoio.

Aos amigos que fiz durante toda minha vida e que mantenho contato até os dias de hoje.

Aos colegas que fiz nesses anos de Instituto Federal.

Ao Prof. Lucas Casanova Silva pela disponibilidade, dedicação e apoio durante os quase 15 meses de construção deste trabalho.

Ao Prof. Marco Aurélio Granero Santos por todas as contribuições dadas à pesquisa.

À todos os professores do Curso de Licenciatura em Matemática, em especial, aos Professores César Adriano Batista e Henrique Marins de Carvalho, exemplos que certamente seguirei durante a carreira.

Aos professores da Escola Estadual Tito Prates da Fonseca, representados pelo diretor, o Prof. Carlos Alberto Vieira, pela contribuição na minha formação e auxílio durante os estágios supervisionados.

Aos professores que tive em toda minha vida, que colaboraram com qualquer tipo de conhecimento acadêmico que obtive.

Àqueles que, direta ou indiretamente, contribuíram para este trabalho.

Por fim, agradeço a Deus pela oportunidade.

"Quanto mais nos elevamos, menores parecemos aos olhos daqueles que não sabem voar"
(Friedrich Wilhelm Nietzsche)

Resumo

Este trabalho tem como objetivo a construção de um modelo realístico da quantidade de usuários da Linha 9 - Esmeralda da CPTM, que passa pelas cidades de São Paulo e Osasco, no período de janeiro de 2008 a dezembro de 2013. Para isto, utilizamos o Método dos Quadrados Mínimos e a ideia da série de Fourier, que consiste em representar de uma função através de uma soma de senos e cossenos. A criação de um modelo é importante, pois auxilia na tomada de decisões, inclusive na cidade de São Paulo, com mais de trezentos quilômetros de sistema metroferroviário. Por estar restrita a um trecho da malha ferroviária, consideramos que esta pesquisa é o passo inicial para o desenvolvimento de modelos maiores e mais complexos.

Palavras-chave: Modelagem. CPTM. Método dos Mínimos Quadrados. Série de Fourier.

Abstract

This work aims to build a realistic model of the number of users of Line 9 - Emerald (CPTM), passing through the cities of São Paulo and Osasco, from January 2008 to December 2013. For this, we used the Method of Least Squares and the idea of Fourier series, which is to represent a function by a sum of sines and cosines. The creation of a model is important because it helps in decision making, including in the city of São Paulo, with more than three hundred kilometers of subway-railroad system. Being restricted to a stretch of the railway network, we believe that this research is the first step towards the development of larger and more complex models .

Keywords: Method of Least Squares. Fourier Series. CPTM. Modeling.

Lista de ilustrações

Lista de tabelas

Lista de abreviaturas e siglas

CPTM	Companhia Paulista de Trens Metropolitanos
SPTRANS	São Paulo Transporte
METRÔ	Companhia do Metropolitano de São Paulo
LSQ	Least Squares

Lista de símbolos

$\in$	Pertence
$\sum$	Somatório
$\prod$	Produtório
ϵ	Epsilon
$\mathbb{R}$	Conjunto dos Número Reais
$\mathbb{R}_+$	Conjunto dos Número Reais positivos

Sumário

1 Introdução

Para Waldvogel et al. (2014), "o ritmo de crescimento do Município de São Paulo foi intenso durante as oito primeiras décadas do século XX [. . .] e cresceu apenas 0,76%, no último período intercensitário de 2000 a 2010". A população estimada em 2014 era de 11.895.893 habitantes (IBGE, 2014). Com esse crescimento populacional, cresce a demanda por transporte público. Em 2013, o sistema de ônibus transportou quase 3 bilhões de pessoas (SPTRANS, 2015) e o sistema metroferroviário, 2,092 bilhões (METRO, 2014).

As notícias de saturação do transporte público são recorrentes. Bruton (1979 apud MENEZES, 2015, p. 14) defende que "os problemas associados aos sistemas de transportes são públicos e, dado o crescimento rápido e constante da população urbana e o explosivo aumento na utilização de veículos motorizados, é natural que tais questões tomem dimensões ainda maiores". Explosivo, no trecho anterior, faz referência ao acelerado aumento no uso de veículos automotores.

Com o crescimento da frota automotiva, o espaço tornou-se cada vez mais disputado. Quintella (2009, p. 20) defende que

> os trens urbanos e metrôs são altamente lucrativos [. . .] no sentido sócio-econômico-ambiental. O transporte público sobre trilhos produz [. . .] um imensurável lucro humanístico e ambiental de grande percepção e possível identificação.

Em relação ao sistema metroferroviário, no ano de 2016, São Paulo conta 74,8 quilômetros de metrô, em seis linhas, sendo cinco gerenciadas pela Companhia do Metropolitano de São Paulo e uma linha pela ViaQuatro (empresa privada); e com os 260,7 quilômetros de trens metropolitanos, em oito linhas, todas gerenciadas pela CPTM. Mas é muito pouco para uma cidade com tantos usuários (NOBRE; BIANCHI, 2014). Para efeitos de comparação, usaremos a quantidade de quilômetros de metrô por milhão de habitantes. Segundo Nobre e Bianchi (2014, p. 21),

> em cidades que apresentam a maior relação, como Londres, Madri, Berlim, Seoul, esse número se encontra entre 30 e 45 quilômetros por milhão de habitantes. São Paulo está [...] numa proporção próxima de 4 quilômetros por milhão de habitantes.

Existe, portanto, a necessidade de expansão do sistema, uma vez que grandes metrópoles mundiais possuem quase 10 vezes mais quilômetros de metrô por milhão de habitantes quando comparadas a São Paulo.

Para Biembengut (1990 apud VIECILI, 2013, p. 29),

> a criação de modelos para interpretar os fenômenos naturais e sociais é inerente ao ser humano. A própria noção de modelo está presente em quase todas as áreas: Arte, Moda, Arquitetura, História, Economia, Literatura, Matemática. Nesse sentido, pode-se dizer que Modelagem Matemática é o processo que descreve um fenômeno para melhor compreendê-lo e estudá-lo, refletindo sobre ele, a fim de obter um modelo matemático.

Desta forma, a criação de um modelo matemático servirá para descrever e refletir sobre os problemas relacionados aos sistemas de transportes públicos, em especial, ao sistema ferroviário, e assim, poderá auxiliar na tomada de decisões. Segundo Sossae, Allevato e Raimundo (2013, p. 21) "uma modelagem eficiente permite fazer previsões, tomar decisões, explicar e entender; enfim, participar do mundo real com capacidade de influenciar em suas mudanças".

Este trabalho propõe a criação de um modelo realístico que avalie a quantidade de usuários da CPTM, na Linha 9-Esmeralda, no período compreendido entre janeiro de 2008 e dezembro de 2013.

De acordo com Junqueira et al. (2013), "[...] para melhorar a qualidade dos serviços prestados [...] há uma necessidade de utilizar e/ou criar ferramentas que levem a uma maior organização e planejamento, auxiliando nas tomadas de decisões".

Optaremos pela restrição à Linha 9 – Esmeralda da CPTM, para evitar que o modelo se torne muito avançado, o que não é nosso objetivo. Aplicá-lo em todas as linhas da CPTM seria um segundo passo depois deste trabalho.

A maioria das estações está situada ao longo de uma importante via da cidade, a Marginal Pinheiros, que possui no seu entorno regiões empresariais e residenciais. Deste modo, acreditamos que este trecho seja mais representativo do que outros.

A metodologia do presente trabalho consiste em fazer pesquisas bibliográficas referentes ao tema proposto, e ainda, uma pesquisa documental, que recorre a "fontes mais diversificadas e dispersas, sem tratamento analítico, tais como: tabelas estatísticas, jornais, revistas, relatórios, documentos oficiais [...]" (GERHARDT; SILVEIRA, 2013). Utilizaremos o Método dos Quadrados Mínimos para modelar os dados fornecidos pela CPTM e a Série de Fourier por causa do comportamento oscilatório dos dados.

A pesquisa está organizada em dois capítulos, além desta Introdução e das Considerações Finais. No Capítulo 2, apresentaremos as bases matemáticas deste trabalho: o Método dos Quadrados Mínimos, as Séries de Fourier e a Interpolação. No Capítulo 3, apresentaremos os dados da pesquisa, faremos a construção do modelo matemático proposto e listaremos as tentativas realizadas até a construção deste.

2 Fundamentos

Neste capítulo, apresentaremos as bases matemáticas dos modelos construídos no Capítulo 3.

Mostraremos definições sobre o Método dos Mínimos Quadrados, o ajuste de curvas através do Método dos Mínimos Quadrados (caso discreto) e sobre as Séries de Fourier. Elas foram utilizadas em todos os modelos, exceto no Modelo 1.

Também apresentaremos definições sobre a Interpolação, em especial, a Forma de Newton, que foi base do Modelo 1.

2.1 Método dos Quadrados Mínimos

As informações desta seção encontram-se na obra "Um curso de Cálculo, volume 2", de Hamilton Luiz Guidorizzi.

2.1.1 Solução LSQ de um sistema linear com uma incógnita

LSQ significa "Least Squares", em português, Mínimos Quadrados. Portanto, a solução LSQ é chamada de solução dos mínimos quadrados. Poderíamos alterar o termo para "Solução MQ", mas optamos por deixar da forma que foi extraída da fonte.

Consideremos um sistema linear S, no plano, com uma incógnita

$$S : \begin{cases} a_{11}t = b_1 \\ a_{21}t = b_2 \end{cases} .$$

Esse sistema, no sentido habitual, poderá ter solução ou não. Terá solução se o ponto $B = (b_1, b_2)$ pertencer à reta r, dada, em forma paramétrica, por

$$r : \begin{cases} x = a_{11}t \\ y = a_{21}t \end{cases} .$$

Se o ponto $B = (b_1, b_2)$ não pertencer à reta r, o sistema não admitirá solução, no sentido habitual, mas admitirá solução LSQ ou solução dos mínimos quadrados.

Dizemos que t_0 é uma solução LSQ, ou seja, a solução dos mínimos quadrados do sistema linear S se $t = t_0$ tornar mínima a distância do ponto $B = (b_1, b_2)$ ao ponto $X = (a_{11}t, a_{21}t)$, sendo t real. Veja a Figura 1:

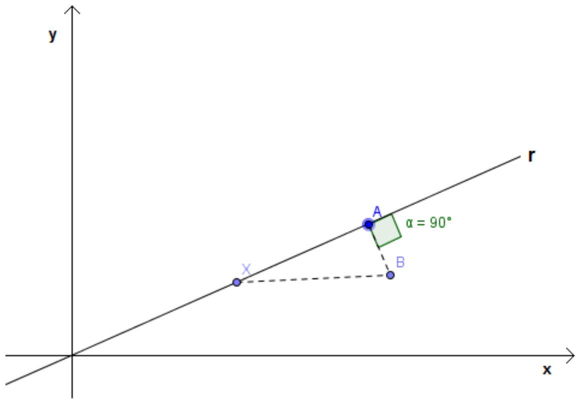

Figura 1 – Método dos Mínimos Quadrados

2.1.2 Solução LSQ de um sistema linear, com uma incógnita, no $\mathbb{R}^n$

Determina-se t que torna mínimo o quadrado da distância euclidiana do ponto $B = (b_1, b_2, \ldots, b_n)$ ao ponto $X = (a_{11}t, a_{21}t, \ldots, a_{n1}t)$. Indicando por W o quadrado da distância de B a X, temos

$$W = \sum_{k=1}^{n} (b_k - ta_{k1})^2.$$

Precisamos determinar os pontos críticos de W, para isso, derivamos a expressão e igualamos a zero.

Derivando, obtemos

$$\frac{dW}{dt} = \sum_{k=1}^{n} 2(b_k - ta_{k1})(-a_{k1}) = 2\sum_{k=1}^{n} ta_{k1}a_{k1} - 2\sum_{k=1}^{n} b_k a_{k1}.$$

Igualando a zero, isso resulta em

$$t = \frac{\sum\limits_{k=1}^{n} b_k a_{k1}}{\sum\limits_{k=1}^{n} a_{k1}a_{k1}}.$$

Como o gráfico de $W = W(t)$ é uma parábola com concavidade voltada para cima, segue que o valor de t acima torna mínimo o valor de W.

2.1.3 Solução LSQ de um sistema linear com mais de uma incógnita, no $\mathbb{R}^n$

Consideremos um sistema com duas incógnitas

$$S : \begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \\ \vdots \\ a_{n1}x + a_{n2}y = b_n \end{cases}.$$

Dizemos que (x_0, y_0) é uma solução LSQ de S se $(x, y) = (x_0, y_0)$ tornar mínima a distância do ponto

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \text{ao ponto } X = \begin{bmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \\ \vdots \\ a_{n1}x + a_{n2}y \end{bmatrix}.$$

Por meio do cálculo diferencial, é possível determinar o ponto que minimiza o quadrado da distância de B a X. Chamando de W o quadrado dessa distância, temos

$$W = \sum_{k=1}^{n} (a_{k1}x + a_{k2}y - b_k)^2.$$

A solução (ou soluções) LSQ de S será (serão) então a (as) solução (soluções) do sistema

$$\begin{cases} \frac{\partial W}{\partial x} = 0 \\ \frac{\partial W}{\partial y} = 0 \end{cases}.$$

Por fim, de $\frac{\partial W}{\partial x} = \sum_{k=1}^{n} 2(a_{k1}x + a_{k2}y - b_k)a_{k1}$ e $\frac{\partial W}{\partial y} = \sum_{k=1}^{n} 2(a_{k1}x + a_{k2}y - b_k)a_{k2}$ resulta

$$\begin{cases} x \sum_{k=1}^{n} {a_{k1}}^2 + y \sum_{k=1}^{n} a_{k1}a_{k2} = \sum_{k=1}^{n} b_k ak1 \\ x \sum_{k=1}^{n} a_{k1}a_{k2} + y \sum_{k=1}^{n} {a_{k2}}^2 = \sum_{k=1}^{n} b_k ak2 \end{cases}.$$

2.1.4 Reta dos Mínimos Quadrados

Seja $\hat{y} = mx + q$ uma reta que queremos determinar. A notação $\hat{y}$ indica que que o valor $\hat{y}$ correspondente ao valor de x é apenas uma estimativa para o verdadeiro valor de y. Para que tal reta passe por todos os pontos, devemos ter

$$\begin{cases} mx_1 + q = y_1 \\ mx_2 + q = y_2 \\ \vdots \\ mx_n + q = y_n \end{cases}.$$

Dizemos que $\hat{y} = mx + q$ é a reta dos mínimos quadrados para os dados da Tabela 1 se (m, q) for a solução LSQ do sistema S.

Tabela 1 – Dados Aleatórios

x	y
x_1	y_1
x_2	y_2
x_3	y_3
$\ldots$	$\ldots$
x_n	y_n

Em geral, o valor $\hat{y}_i$, $\hat{y}_i = mx_i + q$, será apenas uma estimativa para o valor y_i da tabela. Desta forma, cometemos um erro

$$E_i = \hat{y}_i - y_i = mx_i + q - y_i, i = 1, 2, \ldots, n.$$

A soma W dos quadrados dos erros é

$$W = \sum_{i=1}^{n} {E_i}^2 = \sum_{i=1}^{n} (mx_i + q - y_i)^2.$$

Como m e q da reta dos mínimos quadrados $\hat{y}_i = mx + q$ é a solução LSQ do sistema S, resulta que tal reta é determinada de modo que a soma dos quadrados dos erros seja mínima. Logo, a reta dos mínimos quadrados é a reta que minimiza a soma dos quadrados dos erros E_i.

2.2 Ajuste de curvas através do Método dos Mínimos Quadrados

As informações desta seção estão presentes na obra "Cálculo Numérico", de Neide Maria Bertoldi Franco.

2.2.1 Caso discreto

Trataremos da aproximação de uma função $y = f(x)$ por um polinômio de um certo grau m, isto é, $F(x) = P_m(x)$, no caso em que $f(x)$ é dada por pares de pontos, ou seja, caso em que a função é dada por $n+1$ pares de pontos $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$, onde $y_i = f(x_i), i = 0, 1, \ldots, n$, com os $n+1$ pontos $x_0, x_1, \ldots, x_n$ distintos.

Procuramos determinar um polinômio $P_m(x) = a_0 + a_1 x + \ldots + a_m x_m$, de grau no máximo $m, (m < n)$, e tal que $Q = \|f - P_m\|^2$ seja mínimo.

Usando a notação de produto interno

$$< f, g >= \sum_{k=0}^{n} f(x_k) \cdot g(x_k),$$

obtemos

$$Q = \|f - P_m\|^2 =< f - P_m, f - P_m >$$

$$= \sum_{k=0}^{n} [f(x_k) - P_m(x_k)]^2 = \sum_{k=0}^{n} (y_k - P_m(x_k))^2$$

$$= \sum_{k=0}^{n} (y_k - (a_0 + a_1 x_k + \ldots + a_m {x_k}^m))^2$$

.

Assim, dados os $n+1$ pontos distintos $x_0, x_1, \ldots, x_n$ e $n+1$ valores de uma função $y = f(x)$ sobre os pontos x_k, desejamos determinar um polinômio de grau no máximo m menor do que n tal que a soma dos quadrados dos desvios $y_k - P_m(x_k)$ entre os valores de $f(x)$ e $P_m(x)$ calculados nos pontos x_k, seja a menor possível.

Na verdade, precisamos determinar, na classe de todos os polinômios de grau $\leq m$, aquele que minimize Q. Determinemos os coeficientes $a_0, a_1, \ldots, a_m$ de $P_m(x)$. Por definição

$$y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad p = \begin{pmatrix} P_m(x_0) \\ P_m(x_1) \\ \vdots \\ P_m(x_n) \end{pmatrix},$$

onde y e p são vetores do $\mathbb{R}^{n+1}$.

Mas p pode ser escrito como

$$p = a_0 \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + a_1 \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} + a_2 \begin{pmatrix} {x_0}^2 \\ {x_1}^2 \\ \vdots \\ {x_n}^2 \end{pmatrix} + \ldots + a_n \begin{pmatrix} {x_0}^m \\ {x_1}^m \\ \vdots \\ {x_n}^m \end{pmatrix}.$$

Denotando por

$$u_0 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}; \quad u_i = \begin{pmatrix} {x_0}^i \\ {x_1}^i \\ \vdots \\ {x_n}^i \end{pmatrix}, \quad i = 1, 2, \ldots, m,$$

podemos escrever

$$p = a_0 u_0 + a_1 u_1 + \ldots + a_m u_m.$$

Mostremos agora que se os $n+1$ pontos são distintos, então os $m+1$ vetores $u_0, u_1, \ldots, u_m$ são linearmente independentes.

Observemos que p também pode ser escrito como

$$p = \begin{pmatrix} 1 & x_0 & {x_0}^2 & \dots & {x_0}^m \\ 1 & x_1 & {x_1}^2 & \dots & {x_1}^m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & {x_m}^2 & \dots & {x_m}^m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & {x_n}^2 & \dots & {x_n}^2 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}.$$

Seja A a matriz dos coeficientes

$$A = \begin{pmatrix} 1 & x_0 & {x_0}^2 & \dots & {x_0}^m \\ 1 & x_1 & {x_1}^2 & \dots & {x_1}^m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & {x_m}^2 & \dots & {x_m}^m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & {x_n}^2 & \dots & {x_n}^2 \end{pmatrix}.$$

A matriz A possui $n+1$ linhas por $m+1$ colunas, com $n > m$. Seja A' a submatriz quadrada constituída das $m+1$ primeiras linhas e $m+1$ primeiras colunas de A. Assim

$$A' = \begin{pmatrix} 1 & x_0 & {x_0}^2 & \dots & {x_0}^m \\ 1 & x_1 & {x_1}^2 & \dots & {x_1}^m \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & {x_m}^2 & \dots & {x_m}^m \end{pmatrix}.$$

A matriz A' é tal que $detA' = \prod_{i>j}(x_i - x_j)$. Desde que os pontos $x_0, x_1, \dots, x_n$ são distintos, segue que $detA' \neq 0$. Existe, portanto, uma submatriz de A, de ordem $m+1$, que não é singular. Assim, os vetores $u_0, u_1, \dots, u_m$ são linearmente independentes.

Portanto, $u_0, u_1, \dots, u_m$ geram em $\mathbb{R}^{n+1}$ um subespaço vetorial V de dimensão $m+1 < n+1$.

Temos que $y \in \mathbb{R}^{n+1}$ e $p \in V \subset \mathbb{R}^{n+1}$ e queremos que a distância de y a p seja mínima. Isto ocorrerá quando p for a projeção ortogonal de y sobre V.

Os coeficientes $a_0, a_1, \dots, a_m$ do polinômio procurado são então dados pelo sistema normal:

$$\begin{pmatrix} <u_0, u_0> & <u_1, u_0> & \dots & <u_m, u_0> \\ <u_0, u_1> & <u_1, u_1> & \dots & <u_m, u_1> \\ \vdots & \vdots & \vdots & \vdots \\ <u_0, u_m> & <u_1, u_m> & \dots & <u_m, u_m> \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} <y, u_0> \\ <y, u_1> \\ \vdots \\ <y, u_m> \end{pmatrix}.$$

A menos que seja sugerido o produto interno a ser utilizado, usa-se o produto interno usual do $\mathbb{R}^{n+1}$, isto é,

$$< x, y >= \sum_{i=1}^{n} x_i y_i,$$

em que $x = (x_0, x_1, \ldots, x_n)^t$ e $y = (y_0, y_1, \ldots, y_n)^t$.

Observamos que a autora da obra utilizada nesta seção toma (x, y) como notação de produto interno, no entanto, preferimos adotar a notação $< x, y >$ para evitar a confusão com a notação de pares ordenados.

2.3 Séries de Fourier

As informações desta seção encontram-se na obra "Análise de Fourier e equações diferenciais parciais", de Djairo Guedes de Figueiredo.

2.3.1 Funções Periódicas

Uma função $f : \mathbb{R} \to \mathbb{R}$ é periódica de período T se $f(x+T) = f(x)$ para todo x. A função $sen(x)$, por exemplo, é periódica de período 2π.

Em geral, kT é um período, em que k é um inteiro positivo, negativo ou nulo. Se $k = 0$, isso implica em dizer que 0 é um período da função, mas 0 é período de qualquer função. Por isso, sempre consideraremos $T \neq 0$. O menor período positivo é chamado período fundamental.

2.3.2 Convergência uniforme

Uma série numérica $\sum\limits_{n=1}^{\infty} a_n$ converge se a sucessão das reduzidas, também chamadas de somas parciais, converge.

Uma série de funções $\sum\limits_{n=1}^{\infty} u_n x$, em que $u_n : I \to \mathbb{R}$ são funções reais definidas em um subconjunto I de $\mathbb{R}$, convergirá pontualmente se, para cada $x_0 \in I$ fixado, a série numérica $\sum\limits_{n=1}^{\infty} u_n(x_0)$ convergir. Isso equivale a dizer que, dados $\varepsilon > 0$ e $x_0 \in I$, existe um inteiro N, dependendo de ε e de x_0, tal que

$$|\sum_{j=n}^{m} u_j(x_0)| < \varepsilon$$

para todos $n < m$, tais que $n \geq N$.

Uma série de funções $\sum\limits_{n=1}^{\infty} u_n x$ convergirá uniformemente, se, dado $\varepsilon > 0$, existir um inteiro N, dependendo apenas de ε (e não de x), tal que $|\sum\limits_{j=n}^{m} u_j x| < \varepsilon$, para todos $m > n \geq N$.

2.3.2.1 Teste M de Weierstrass

Seja $\sum\limits_{n=1}^{\infty} u_n(x)$ uma série de funções $u_n : I \to \mathbb{R}$ definida em um subconjunto I de $\mathbb{R}$. Suponha que existam constantes $M_n \geq 0$ tais que

$$|u_n(x)| \leq M_n, \forall x \in I$$

e que a série numérica $\sum\limits_{n=1}^{\infty} M_n$ convirja. Então, a série de funções $\sum\limits_{n=1}^{\infty} u_n(x)$ converge uniforme e absolutamente em I.

A demonstração deste teorema encontra-se no Anexo A.

2.3.3 Coeficientes de Fourier

Sejam

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)cos\frac{n\pi x}{L}dx, n \geq 0,$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)sen\frac{n\pi x}{L}dx, n \geq 1;$$

e $f : \mathbb{R} \to \mathbb{R}$ uma função periódica de período $2L$, integrável e absolutamente integrável em cada intervalo limitado; em particular, $\int_{-L}^{L} |f(x)|dx < \infty$. Os números a_n e b_n são definidos como coeficientes de Fourier da função f.

2.3.4 Série de Fourier

Dada uma função $\mathbb{R} \to \mathbb{R}$ periódica de período $2L$, integrável e absolutamente integrável, podemos escrever

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n cos\frac{n\pi x}{L} + b_n sen\frac{n\pi x}{L}),$$

e isso significa que a expressão do lado direito é a série de Fourier.

2.4 Interpolação

As informações desta seção encontram-se na obra "Cálculo Numérico, Aspectos Teóricos e Computacionais", de Márcia A. Gomes Ruggiero e Vera Lúcia da Rocha Lopes.

Interpolar uma função $f(x)$ consiste em aproximar essa função por uma outra função $g(x)$, escolhida entre uma classe de funções definida *a priori* e que satisfaça algumas propriedades. A função $g(x)$ é então usada em substituição à função $f(x)$.

A necessidade de se efetuar esta substituição surge em várias situações, como por exemplo, quando são conhecidos somente os valores numéricos da função para um conjunto de pontos e é necessário calcular o valor da função em um ponto não tabelado.

Tomemos $(n+1)$ pontos distintos, $x_0, x_1, \ldots, x_n$, chamados nós da interpolação, e os valores de $f(x)$ nesses pontos, $f(x_0), f(x_1), \ldots, f(x_n)$. A forma de interpolação de $f(x)$ a seguir consiste em se obter uma determinada função $g(x)$ tal que

$$\begin{cases} g(x_0) = f(x_0) \\ g(x_1) = f(x_1) \\ g(x_2) = f(x_2) \\ \qquad \vdots \\ g(x_n) = f(x_n) \end{cases}.$$

Consideraremos que $g(x)$ pertence à classe das funções polinomiais, mas existem outras formas de interpolação polinomial como, por exemplo, a fórmula de Taylor e a interpolação por polinômios de Hermite.

Assim como $g(x)$ foi escolhida entre as funções polinomiais, $g(x)$ poderia ser uma função trigonométrica, entre outras.

2.4.1 Interpolação Polinomial

Dados os pontos $(x_0, f(x_0)), (x_1, f(x_1)), \ldots, (x_n, f(x_n))$, portanto, $(n+1)$ pontos, queremos aproximar $f(x)$ por um polinômio $p_n(x)$, de grau menor ou igual a n, tal que

$$f(x_k) = p_n(x_k),\ k = 0, 1, 2, \ldots, n.$$

2.4.1.1 Teorema 1

Existe um único polinômio $p_n(x),$ de grau menor ou igual a n, tal que:

$$p_n(x_k) = f(x_k), k = 0, 1, 2, \ldots, n, \text{ desde que } x_k \leq x_j,\ j \leq k.$$

Seja $p_n(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$. Portanto, obter $p_n(x)$ significa obter os coeficientes $a_0, a_1, a_2, \ldots, a_n$.

Da condição $p_n(x_k) = f(x_k), \forall k = 0, 1, 2, \ldots, n$, segue o seguinte sistema linear com $n+1$ equações e $n+1$ variáveis, $a_0, a_1, \ldots, a_n$:

$$\begin{cases} a_0 + a_1x_0 + a_2{x_0}^2 + \ldots + a_n{x_0}^n = f(x_0) \\ a_0 + a_1x_1 + a_2{x_1}^2 + \ldots + a_n{x_0}^n = f(x_1) \\ \qquad\qquad \vdots \\ a_0 + a_1x_n + a_2{x_n}^2 + \ldots + a_n{x_n}^n = f(x_n) \end{cases}$$

A matriz A dos coeficientes é

$$A = \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ . & . & . & & . \\ . & . & . & & . \\ . & . & . & & . \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix}$$

que é uma matriz de Vandermonde e, portanto, desde que $x_0, x_1, \dots, x_n$ sejam pontos distintos, temos $det(A) \neq 0$ e, então, o sistema linear admite solução única.

2.4.2 Forma de Newton

A forma de Newton para o polinômio $p_n(x)$ que interpola $f(x)$ em $x_0, x_1, \dots, x_n$, $(n+1)$ pontos distintos, é

$$p_n(x) = d_0 + d_1(x - x_0) + d_2(x - x_0)(x - x_1) + \dots + d_n(x - x_0)(x - x_1)\dots(x - x_{n-1}).$$

Os coeficientes reais d_k, $k = 0, 1, \dots, n$ são diferenças divididas de ordem k entre os pontos $(x_j, f(x_j))$, $j = 0, 1, \dots, k$.

2.4.2.1 Operador das Diferenças Divididas

Seja $f(x)$ uma função tabelada em $n+1$ pontos distintos: $x_0, x_1, \dots, x_n$. Definimos o operador das diferenças divididas por:

$$\left[\begin{array}{l} f[x_0] = f(x_0) \\ f[x_0, x_1] = \frac{f[x_1]-f[x_0]}{x_1-x_0} = \frac{f(x_1)-f(x_0)}{x_1-x_0} \\ \\ f[x_0, x_1, x_2] = \frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0} \\ \\ f[x_0, x_1, x_2, x_3] = \frac{f[x_1,x_2,x_3]-f[x_0,x_1,x_2]}{x_3-x_0} \\ \\ \qquad\quad \vdots \qquad\qquad\qquad\quad \vdots \\ \\ f[x_0, x_1, x_2, \dots, x_n] = \frac{f[x_1,x_2,\dots,x_n]-f[x_0,x_1,x_2,\dots,x_{n-1}]}{x_n-x_0} \end{array}\right.$$

Dizemos que $f[x_0, x_1, \dots, x_k]$ é a diferença dividida de ordem k da função $f(x)$ sobre os $k+1$ pontos: $x_0, x_1, \dots, x_k$.

Dada uma função $f(x)$ e conhecidos os valores que $f(x)$ assume nos pontos distintos $x_0, x_1, \dots, x_n$, é possível construir a Tabela 2:

Tabela 2 – Tabela das Diferenças Divididas

x	Ordem 0	Ordem 1	Ordem 2	...	Ordem n
x_0	$f[x_0]$				
		$f[x_0, x_1]$			
x_1	$f[x_1]$		$f[x_0, x_1, x_2]$		
		$f[x_1, x_2]$		$\ddots$	
x_2	$f[x_2]$		$f[x_1, x_2, x_3]$		
		$f[x_2, x_3]$		...	$f[x_0, x_1, x_2, \ldots, x_n]$
x_3	$f[x_3]$		$f[x_2, x_3, x_4]$		
		$f[x_3, x_4]$	$\vdots$		
x_4	$f[x_4]$	$\vdots$	$f[x_{n-2}, x_{n-1}, x_n]$		
$\vdots$	$\vdots$	$f[x_{n-1}, x_n]$			
x_n	$f[x_n]$				

Uma propriedade das diferenças divididas é que $f[x_0, x_1, \ldots, x_n]$ é simétrica nos argumentos, ou seja, $f[x_0, x_1, \ldots, x_k] = f[x_{j_0}, x_{j_1}, \ldots, x_{j_k}]$, onde $j_0, j_1, \ldots, j_k$ é qualquer permutação de $0, 1, \ldots, k$.

2.4.2.2 Interpretação da Forma de Newton

A interpretação da Forma de Newton encontra-se nas Notas de Aula de Cálculo Númerico, do Professor José Álvaro Tadeu Ferreira, da Universidade Federal de Ouro Preto.

Dada uma função $y = f(x)$, a sua primeira derivada é definida como

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

Sendo $(x_i, y_i), i = 0, 1, \ldots, n$ um conjunto de pontos da função, então

$$\lim_{h \to 0} \frac{f(x_i+h) - f(x_i)}{h}.$$

Se $x_i + h = x_{i+1}$, então $h = x_{i+1} - x_i$. Logo,

$$f'(x_i) = \lim_{x_i \to x_{i+1}} \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}.$$

Sendo $(x_i, y_i), i = 0, 1, \ldots, n$ uma sequência de pontos de uma função $y = f(x)$, com abscissas distintas, define-se operador da diferença dividida de primeira ordem como

$$Dy_i = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \ i = 0, 1, \ldots, n-1.$$

Este operador, portanto, é uma aproximação do valor numérico da primeira derivada de uma função em um ponto. As diferenças divididas de ordem superior são aproximações para derivadas de ordem superior.

A diferença dividida de segunda ordem é definida como

$$D^2 y_i = \frac{Dy_{i+1} - Dy_i}{x_{i+2} - x_i},\ i = 0, 1, \ldots, n-2.$$

A diferença dividida de terceira ordem é definida como

$$D^3 y_i = \frac{D^2 y_{i+1} - D^2 y_i}{x_{i+3} - x_i},\ i = 0, 1, \ldots, n-3.$$

Generalizando, a diferença dividida de ordem n é definida como

$$D^n y_i = \frac{D^{n-1} y_{i+1} - D^{n-1} y_i}{x_{i+k} - x_i},\ i = 0, 1, \ldots, n-k \text{ e } k = 1, \ldots, n.$$

A diferença dividida de ordem zero é dada por $D^0 y_i = y_i, i = 0, 1, \ldots, n$.

Ferreira (2013) usa uma notação diferente de Ruggiero e Lopes (2012) para definir as diferenças divididas. O primeiro define como $D^i y_0, i = 0, 1, \ldots, n$ e o segundo, como d_k, $k = 0, 1, \ldots, n$. Portanto,

$$d_k = D^i y_0,\ \text{sendo } k = 0, 1, \ldots, n \text{ e } i = 0, 1, \ldots, n.$$

3 Modelos

Neste capítulo, pretendemos encontrar uma função que seja uma boa aproximação para os dados da Tabela 16, disponível no Anexo J, que foi obtida através dos dados das tabelas 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 e 15, que estão nos Anexos D, E, F, G, H e I.

Para isso, usaremos as definições vistas no Capítulo 2.

No mês de setembro de 2015, enviamos ao Departamento de Marketing da CPTM um ofício (disponível no Anexo B) solicitando informações sobre a Linha 9 - Esmeralda.

Os dados foram obtidos, no mês de outubro, junto ao Departamento de Planejamento de Transporte da CPTM, através do senhor Rodrigo Sartoratto de Alencar, chefe do departamento. A companhia cedeu uma planilha eletrônica com os dados catracados por estação, ou seja, o número mensal de usuários que entrou nas estações da Linha 9 - Esmeralda, no período compreendido entre janeiro de 2008 e dezembro de 2013, totalizando 72 meses.

A carta que atesta a entrega dos dados está disponível no Anexo C.

3.1 Erro

Exceção feita à Tentativa 1, nas demais tentativas, no Pré-modelo e no Modelo, o erro médio ($E_{médio}$) foi calculado da seguinte forma:

$$E_{médio} = \frac{\sum_{i=1}^{72} E_i}{72}.$$

É a soma dos erros em cada um dos 72 pontos dividido por 72.

Já o erro em cada ponto foi definido por

$$E_i = \frac{\sqrt{(h(x_i) - f(x_i))^2}}{f(x_i)},$$

sendo $f(x_i)$ a função original e $h(x_i)$ a função do modelo e $i = 1, \ldots, 72$.

3.2 Apresentação dos dados

A partir da Tabela 16, construímos um gráfico (Figura 2) que serviu para analisar os dados e definir as funções "candidatas"a uma boa aproximação:

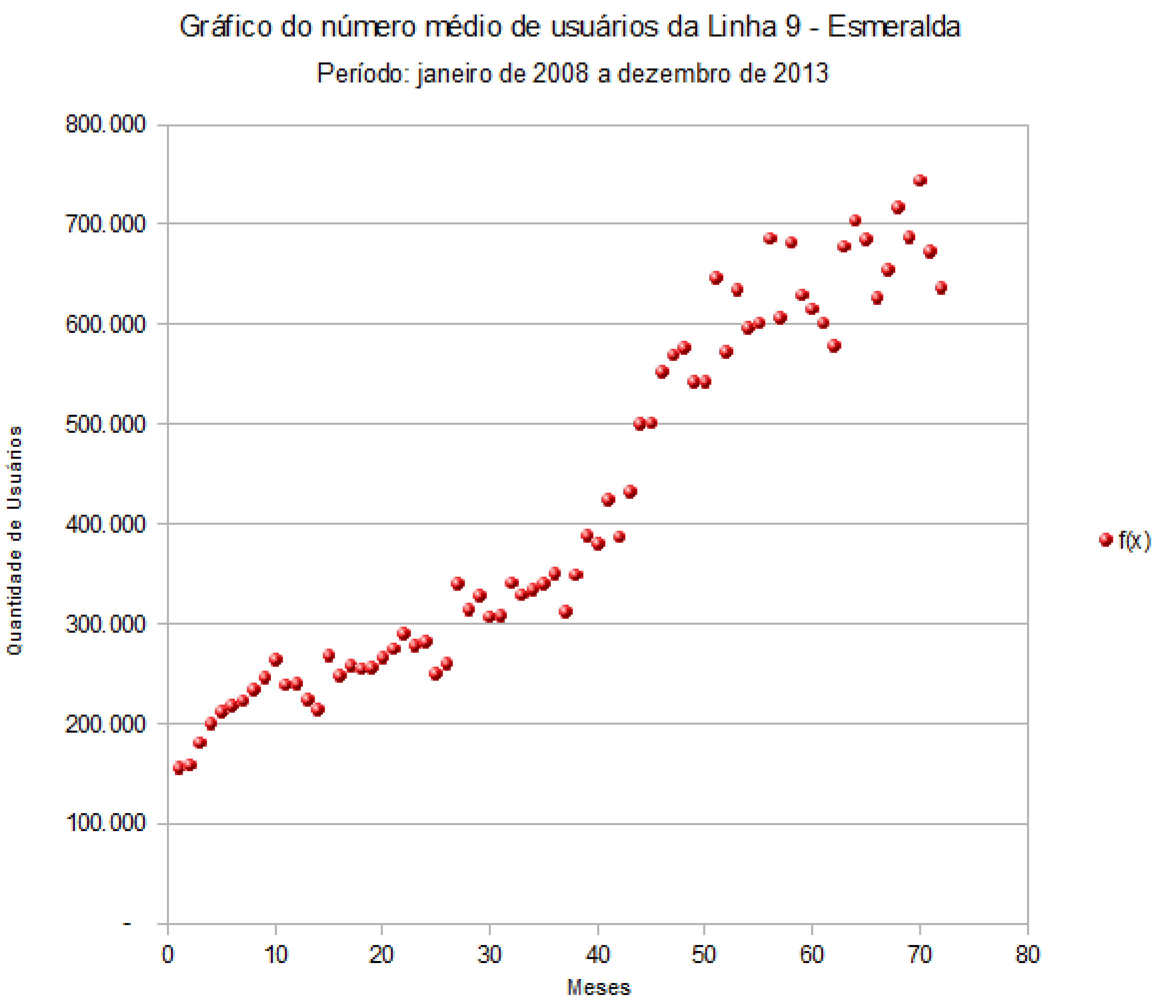

Figura 2 – Gráfico do número de usuários da Linha 9 - Esmeralda no período de janeiro de 2008 a dezembro de 2013.

Os dados informados pela CPTM sugerem crescimento e certa periodicidade. Nas próximas seções, apresentaremos as etapas de construção do Modelo Final (seção 3.10). São os Modelos 1 a 7, descartados, mas que fazem parte da evolução da pesquisa.

No Modelo 1, tentamos aproximar uma função através da interpolação polinomial. Nos Modelos 2 a 7 e no Modelo Final, usamos o ajuste de curvas através do Método dos Mínimos Quadrados. A série de Fourier é usada a partir do Modelo 5.

3.3 Modelo 1

Nosso primeiro modelo envolveu a interpolação polinomial, ou seja, um ajuste de curva no qual a curva ajustada coincide necessariamente com os valores medidos. Testamos a fórmula para seis valores da Estação Osasco, os dados de janeiro a junho do ano de 2008.

Chegamos ao polinômio $h(x) = 386987 - 399390x + 267157x^2 - 79925x^3 + 11121x^4 -$

$586x^5$, definido através da Forma de Newton. Vejamos a Tabela 3:

Tabela 3 – Valores do polinômio interpolador

x	h(x)
1	185364
1,5	171109
2	176619
2,5	186603
3	193658
3,5	196078
4	195651
4,5	195466
5	197712
5,5	201483
6	200579
7	120294
12	-19255981

Notamos que os valores para $x = 1,5$, $x = 2,5$, $x = 3,5$ e $x = 4,5$ (aqui definimos $1,5$, por exemplo, como o número de usuários na metade do mês de fevereiro de 2008, ou seja, a quantidade de usuários de janeiro mais a quantidade de usuários da primeira metade de fevereiro) não são absurdos. No entanto, o valor para $x = 7$ destoa e, para $x = 12$, é impossível. Esta é uma das razões pela qual descartamos o Modelo 1.

A outra razão é que uma interpolação com seis valores gerou um polinômio do quinto grau. Para 72 valores, o polinômio gerado seria, no máximo, do 71º grau, bem trabalhoso. Também não há garantia de que, fora dos pontos x_i, a função interpoladora coincida com a original. Neste caso, a função não coincide e o erro é alto.

Optamos por não encontrar uma função através da interpolação polinomial. Nos próximos modelos, utilizaremos o Método dos Mínimos Quadrados.

3.4 Modelo 2

Neste e nos próximos modelos, faremos um ajuste de curvas através do Método dos Mínimos Quadrados, como visto na seção 2.2. Usaremos uma combinação de funções para fazer uma aproximação da função desejada.

Observando o comportamento crescente e periódico do gráfico (Figura 2), acreditamos que as funções $g_2 = \text{sen}(x)$, periódica e positiva (no primeiro quadrante), e $g_3 = \sqrt{x}$, positiva e crescente em todo o seu domínio, são suficientes para fazer tal aproximação.

Assim, a função do modelo é da forma $h(x) = a_0 + a_1 \cdot \text{sen}(x) + a_2\sqrt{x}$.

Os coeficientes foram determinados pelo sistema

$$\begin{cases} 72a_0 + 1,93a_1 + 411,33a_2 = 30219194,25 \\ 1,93a_0 + 36,11a_1 + 9,21a_2 = 1161746,08 \\ 411,33a_0 + 9,21a_1 + 2628a_2 = 195877844,05 \end{cases} .$$

As soluções deste sistema e dos outros que serão apresentados no trabalho foram feitas através da Regra de Cramer (ver Anexo K).

Chegamos ao modelo definido por

$$h(x) = -58546,81 + 13966,44 \cdot \operatorname{sen}(x) + 83649,65\sqrt{x}.$$

Veja a Figura 3:

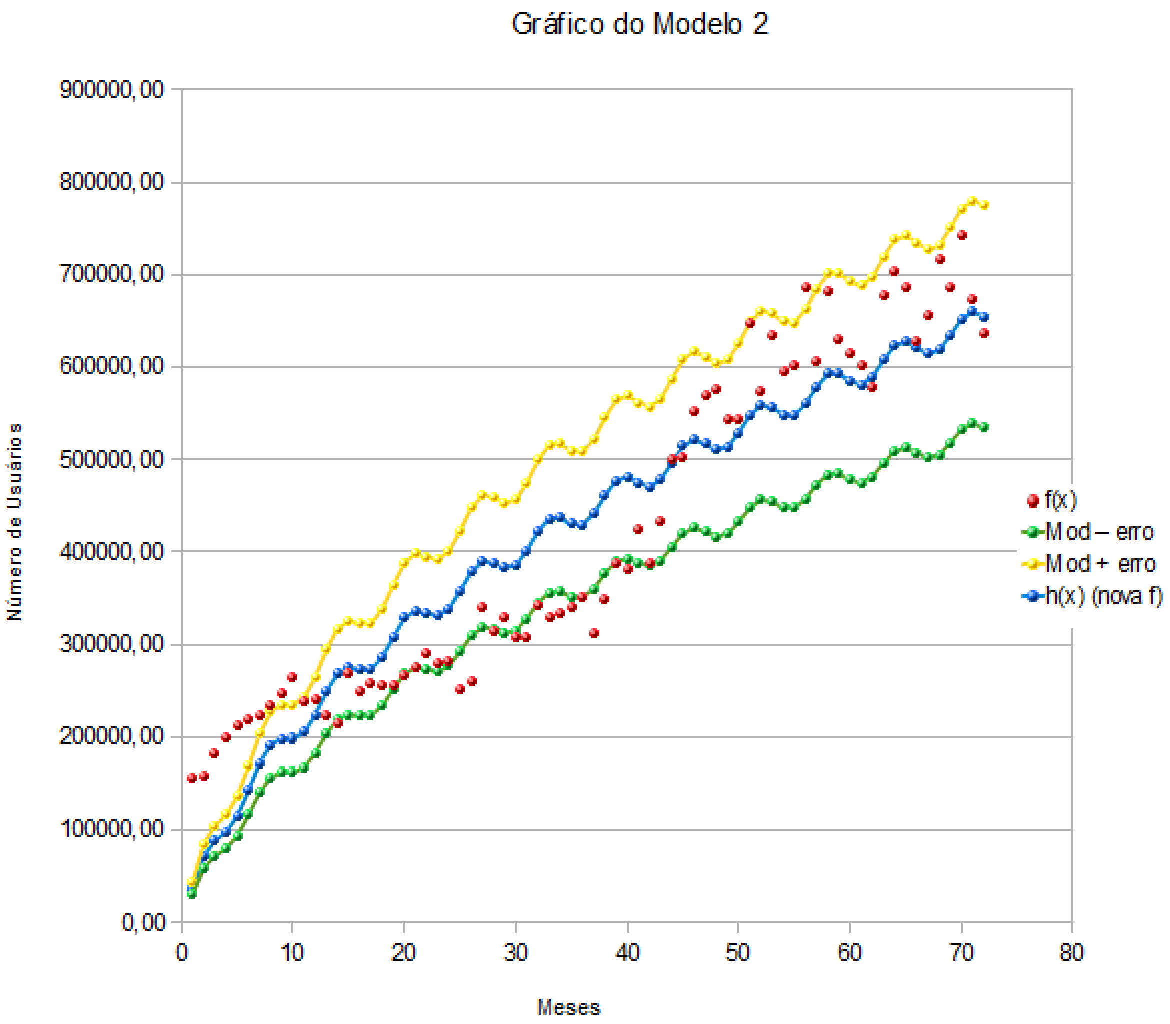

Figura 3 – Gráfico do Modelo 2

O erro médio calculado foi de 18,31%. O gráfico mostra oscilação e crescimento, no entanto, boa parte dos dados está abaixo da função do modelo menos o erro médio. Por essa razão, descartamos este modelo.

No Modelo 3, substituiremos a função $\sqrt{x}$ pela função x^2. Manteremos a função $sen(x)$ por ser crescente no intervalo $[1; 72]$. Lembramos que a função $\cos(x)$ também é periódica, mas é decrescente no intervalo citado.

3.5 Modelo 3

A Figura 4 mostra o gráfico do Modelo 3:

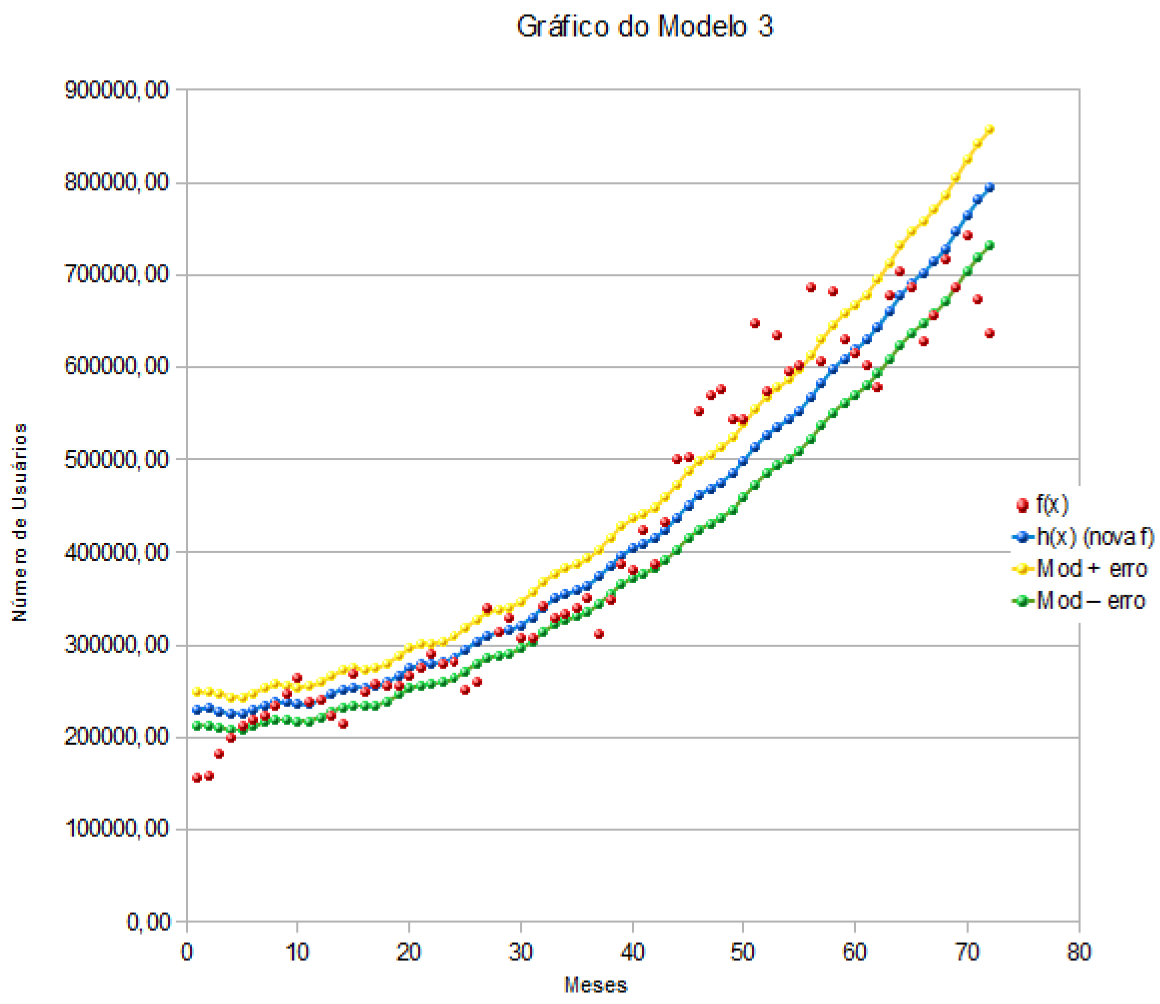

Figura 4 – Gráfico do Modelo 3

A função polinomial do segundo grau tem o comportamento crescente quando $x \in \mathbb{R}_+$. A função deste modelo, portanto, é da forma $h(x) = a_0 + a_1 \cdot \text{sen}(x) + a_2x^2$.

Os coeficientes foram determinados pelo sistema

$$\begin{cases} 72a_0 + 1,93a_1 + 127020a_2 = 30219194,25 \\ 1,93a_0 + 36,11a_1 + 5282,99a_2 = 1161746,08 \\ 127020a_0 + 5282,99a_1 + 400544868a_2 = 72579678693,46 \end{cases}.$$

O modelo é $h(x) = 227046,3 + 4068,62 \cdot \text{sen}(x) + 109,15x^2$.

Notemos que no Modelo 2, o comportamento da função foi mais oscilatório que o do Modelo 3. Por outro lado, o crescimento do último foi maior.

O erro médio calculado foi de $7,89\%$. Optamos por descartar este modelo pois, para valores entre os pontos $h(50)$ e $h(64)$, por exemplo, os valores da função original são superiores aos do modelo mais o erro.

No próximo modelo, incluiremos a função x na combinação para tentar corrigir o problema.

3.6 Modelo 4

Neste ensaio, faremos uma combinação entre a função x^2, a função x e a função $sen(x)$. Acreditamos que a inclusão da função do primeiro grau diminuirá o erro médio.

A seguir, o gráfico do modelo (Figura 5):

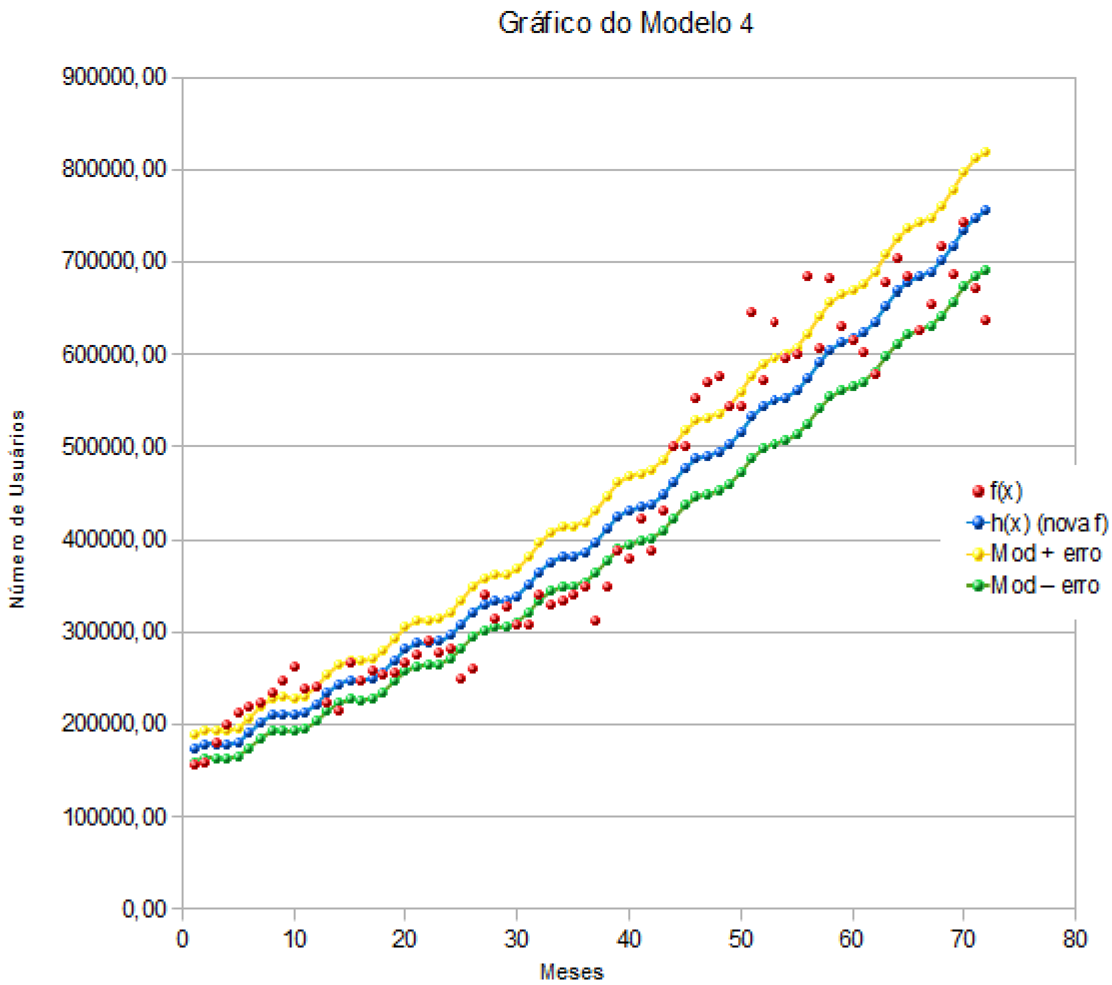

Figura 5 – Gráfico do Modelo 4

A função do modelo é da forma $h(x) = a_0 + a_1 \cdot \text{sen}(x) + a_2 x + a_3 x^2$.

Os coeficientes foram determinados pelo sistema

$$\begin{cases} 72a_0 + 1,93a_1 + 2628a_2 + 127020a_3 = 30219194,25 \\ 1,93a_0 + 36,11a_1 + 73,15a_2 + 5282,99 = 1161746,08 \\ 2628a_0 + 73,15a_1 + 127020a_2 + 6906384a_3 = 1359430269,59 \\ 127020a_0 + 5282,99a_1 + 6906384a_2 + 400544868a_3 = 72579678693,46 \end{cases}.$$

O modelo calculado foi

$$h(x) = 163611,83 + 6803,75 \cdot \text{sen}(x) + 4593,75x + 50,02x^2.$$

A inclusão da função x não diminuiu o erro médio, ao contrário, aumentou para $8,49\%$, razão pela qual descartamos o modelo.

Nos próximos modelos, passaremos a utilizar a definição de Série de Fourier para a construção dos modelos, ou seja, uma soma de senos e cossenos (no caso, uma soma finita).

3.7 Modelo 5

Neste modelo, tentamos reproduzir a ideia da Série de Fourier. Para isso, usamos uma soma de senos e cossenos em sua construção.

Observando a Figura 2, determinamos que o período da função tem 12 meses. Da Série de Fourier, se o período é $2L$, logo $2L = 12$, e então, $L = 6$.

Assim, chegamos a $cos(\frac{\pi}{6}x)$ e $sen(\frac{\pi}{6}x)$.

A função do modelo é da forma $h(x) = a_0 + a_1 \cdot \text{sen}(\frac{\pi}{6}x) + a_2 \cdot \cos(\frac{\pi}{6}x) + a_3 \cdot \text{sen}(\frac{2\pi}{6}x) + a_4 \cdot \cos(\frac{2\pi}{6}x) + a_5 \cdot \text{sen}(\frac{3\pi}{6}x) + a_6 \cdot \cos(\frac{3\pi}{6}x)$.

Os coeficientes foram determinados pelo sistema

$$\begin{cases} 72a_0 = 30219194,25 \\ 36a_1 = -1373784,53 \\ 36a_2 = 35390,0652 \\ 36a_3 = -989914,62 \\ 36a_4 = -204072,59 \\ 36a_5 = -430472 \\ 36a_6 = -501642,27 \end{cases}.$$

O modelo calculado foi $h(x) = 419711 - 38160,68 \cdot \text{sen}(\frac{\pi}{6}x) + 983,06 \cdot \cos(\frac{\pi}{6}x) - 27497,63 \cdot \text{sen}(\frac{2\pi}{6}x) - 5668,68 \cdot \cos(\frac{2\pi}{6}x) - 11957,56 \cdot \text{sen}(\frac{3\pi}{6}x) + 13934,51 \cdot \cos(\frac{3\pi}{6}x)$.

Veja o gráfico do Modelo 5 (Figura 6):

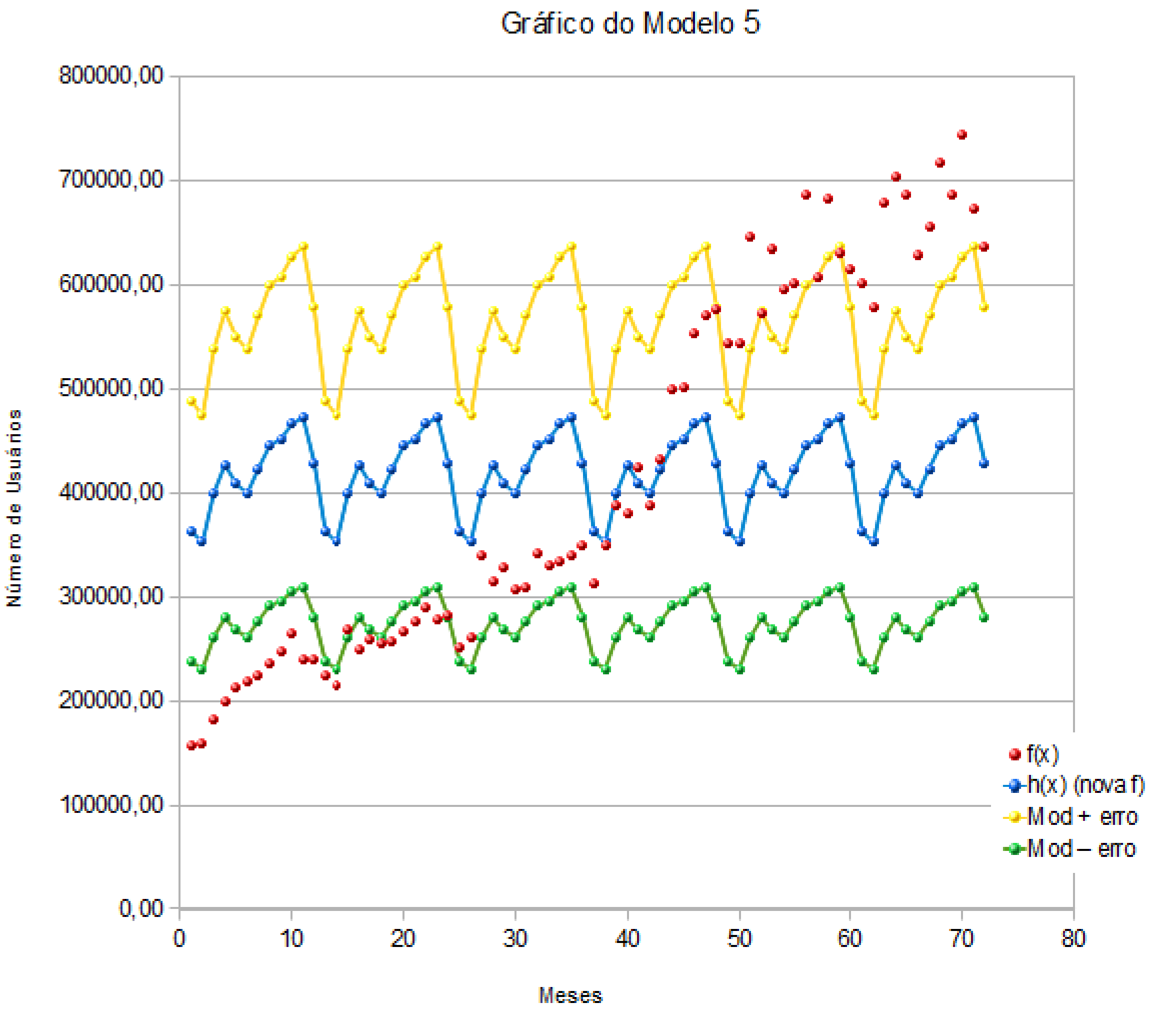

Figura 6 – Gráfico do Modelo 5

Percebemos que há oscilação e há uma "onda"a cada período de 12 meses. No entanto, o gráfico de $h(x)$ não "sobe", ou seja, não há um crescimento de acordo com $f(x)$.

Além disso, o erro é $34,55\%$, o maior até aqui, o que descartou esta possibilidade.

No Modelo 6, faremos uma combinação com $sen(x), sen(2x), cos(x), cos(2x)$ e a função x, que tem o comportamento crescente. A intenção é que a função oscile e cresça.

3.8 Modelo 6

Neste modelo, combinaremos as funções seno, cosseno e a função x. Aqui desprezamos o fato de $L = 6$, como definido para o modelo anterior.

A função é da forma $h(x) = a_0 + a_1 \cdot \text{sen}(x) + a_2 \cdot \cos(x) + a_3 \cdot \text{sen}(2x) + a_4 \cdot \cos(2x) + a_5 x$.

Os coeficientes foram determinados pelo sistema

$$\begin{cases} 72a_0 + 1,93a_1 - 0,75a_2 - 0,2a_3 - 0,22a_4 + 2628a_5 = 30219194,25 \\ 1,93a_0 + 36,11a_1 - 0,1a_2 + 0,04a_3 - 0,76a_4 + 73,15a_5 = 1161746,08 \\ -0,75a_0 - 0,10a_1 + 35,89a_2 + 1,17a_3 - 0,79a_4 - 20,23a_5 = -304640,72 \\ -0,2a_0 + 0,04a_1 + 1,17a_2 + 36,02a_3 - 0,27a_4 - 37,99a_5 = -204733,16 \\ -0,22a_0 - 0,76a_1 - 0,79a_2 - 0,27a_3 + 35,98a_4 + 19,97a_5 = 79608,63 \\ 2628a_0 + 73,15a_1 - 20,23a_2 - 37,99a_3 + 19,97a_4 + 127020a_5 = 1359430269,59 \end{cases}.$$

O modelo, portanto:

$$h(x) = 118318,84 + 9096,43 \cdot \operatorname{sen}(x) - 1492,58 \cdot \cos(x) + 3702,16 \cdot \operatorname{sen}(2x)$$
$$-1456,02 \cdot \cos(2x) + 8250,37x.$$

A Figura 7 mostra o gráfico:

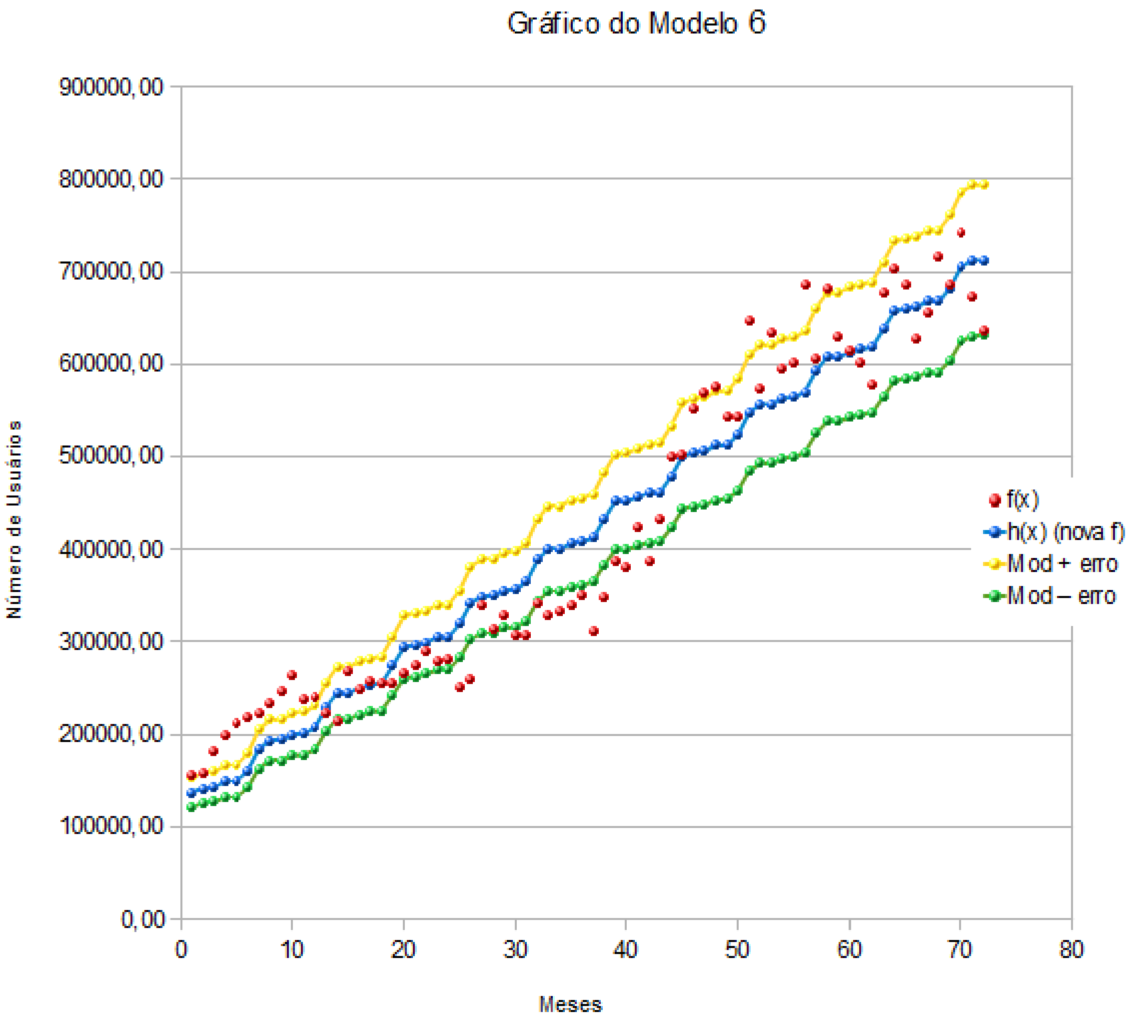

Figura 7 – Gráfico do Modelo 6

Há oscilação e há crescimento, mas o erro médio calculado é de 11,47%, maior que o Modelo 3. Portanto, foi descartado.

3.9 Modelo 7

No Modelo 7, retomaremos a ideia do Modelo 5, mas com algumas alterações: excluiremos a função $cos(\frac{3\pi}{6}x)$ e adicionaremos as funções x e x^2. Acreditamos que o modelo irá oscilar, como no Modelo 5, e crescer, a exemplo do Modelo 4.

A função do modelo é da forma $h(x) = a_0 + a_1 \cdot \text{sen}(\frac{\pi}{6}x) + a_2 \cdot \cos(\frac{\pi}{6}x) + a_3 \cdot \text{sen}(\frac{2\pi}{6}x) + a_4 \cdot \cos(\frac{2\pi}{6}x) + a_5 \cdot \text{sen}(\frac{3\pi}{6}x) + a_6 x + a_7 x^2$.

Veja o gráfico do modelo (Figura 8):

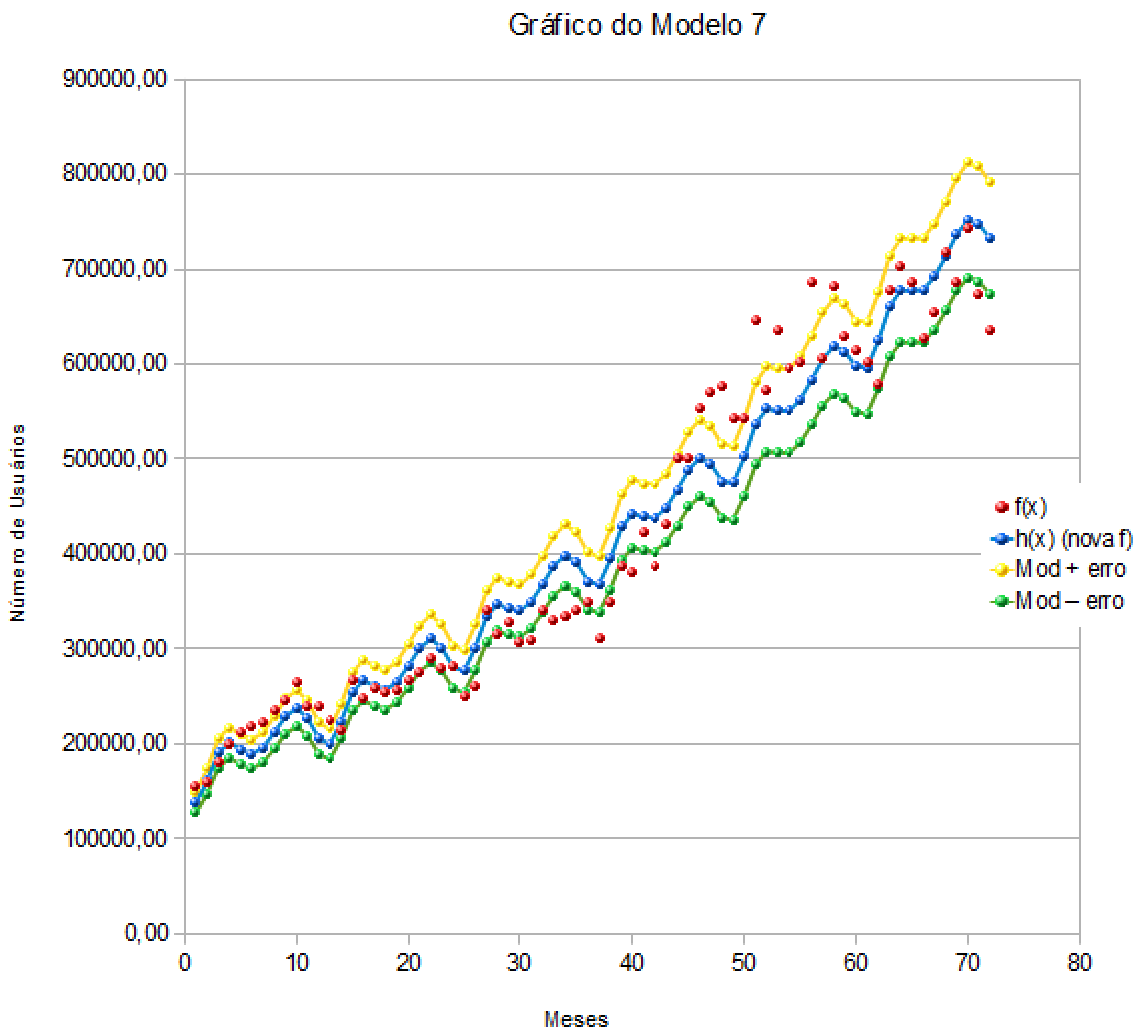

Figura 8 – Gráfico do Modelo 7

Os coeficientes foram determinados pelo sistema

$$\begin{cases} 72a_0 + 127015a_7 = 30219194,25 \\ 36a_1 - 9677,4398a_7 = -1373784,5344 \\ 36a_2 + 3126,55a_7 = 35390,07 \\ 36a_3 - 4493,81a_7 = -989914,6219 \\ 36a_4 + 2737,5a_7 = -204072,5939 \\ 36a_5 - 2593a_7 = -430472,0026 \\ 36a_6 + 2668a_7 = 501642,2767 \\ \\ 127015a_0 - 9677,4398a_1 + 3126,54a_2 - 4493,81a_3 + \\ 2737,5a_4 - 2593a_5 + 2668a_6 + 400544851a_7 = 72579678693,4582 \end{cases} .$$

A função calculada é $h(x) = 163611,83 - 7721,71 \cdot \text{sen}(\frac{\pi}{6}x)$ $-7941,79 \cdot \cos(\frac{\pi}{6}x) - 13367,27 \cdot \text{sen}(\frac{2\pi}{6}x)$ $-14032,92 \cdot \cos(\frac{2\pi}{6}x) - 3801,55 \cdot \text{sen}(\frac{3\pi}{6}x)$ $+4419,51x + 51,88x^2$.

O erro médio calculado foi de $7,25\%$, o menor até agora. No entanto, alguns detalhes precisam ser observados. O modelo se ajusta bem ãté o mês 44. Daí em diante, pontos de $f(x)$ superam a função do modelo mais o erro.

Decidimos "quebrá-lo"no mês 44 e a decisão está baseada no crescimento do número de usuários da Estação Pinheiros, no mês de junho de 2011 (ver Tabelas 10 e 11). Descreveremos esta situação no Modelo Final.

3.10 Modelo Final

A integração da Linha 9 - Esmeralda com a Linha 4 - Amarela, na Estação Pinheiros, a partir de junho de 2011, impactou o número médio de usuários do sistema, quando, na Estação Pinheiros, passou de 385.112 (em junho) para 909.937 (em julho), um aumento de 136%.

A Estação Pinheiros da Linha 4 - Amarela foi inaugurada em 16 de maio de 2011. O número de usuários passou de 191.610 em abril (antes da inauguração) para 385.112, em junho (após a inaguração), aumento de 101%.

Percebemos que, de certa forma, a integração prejudicou os modelos criados até agora. Decidimos dividir o Modelo 7 em duas partes. Inicialmente, a quebra foi feita no mês 42 (junho de 2011), o erro para o intervalo $[1;43)$ foi de $3,42\%$ e no intervalo $[43;72]$, $5,5\%$. Logo, o erro médio para o intervalo todo, fazendo uma média ponderada, foi de $4,29\%$.

Mas, ao dividir o Modelo 7 no mês 43, para o intervalo $[1;44)$, o erro foi de $4,13\%$ e para o intervalo $[44;72]$, de $4,47\%$. O erro para o intervalo $[1;72]$ foi de $4,27\%$, menor que o da divisão anterior, que foi descartada.

Os coeficientes foram obtidos de forma parecida com a do Modelo 7, no entanto, por demandar mais cálculos (note que no sistema que determinou os coeficientes do Modelo, muitos produtos internos zeraram), não colocaremos os sistemas neste texto.

$$h(x) = \begin{cases} 191907,20 - 8742,27 \cdot \text{sen}(\frac{\pi}{6}x) - 9819,12 \cdot \cos(\frac{\pi}{6}x) - 12328,5 \cdot \text{sen}(\frac{2\pi}{6}x) - \\ 5427,02 \cdot \cos(\frac{2\pi}{6}x) - 5439,19 \cdot \text{sen}(\frac{3\pi}{6}x) + 2726,18x + 49,81x^2, \\ \text{se } 1 \leq x < 44, \\ \\ -492792,03 - 8942,09 \cdot \text{sen}(\frac{\pi}{6}x) - 7997,08 \cdot \cos(\frac{\pi}{6}x) - 22413,61 \cdot \text{sen}(\frac{2\pi}{6}x) \\ -13569,34 \cdot \cos(\frac{2\pi}{6}x) + 18377,45 \cdot \text{sen}(\frac{3\pi}{6}x) + 33020,33x - 233,41x^2, \\ \text{se } 44 \leq x \leq 72. \end{cases}$$

A Figura 9 mostra o gráfico do Modelo Final para o intervalo $[1;44)$:

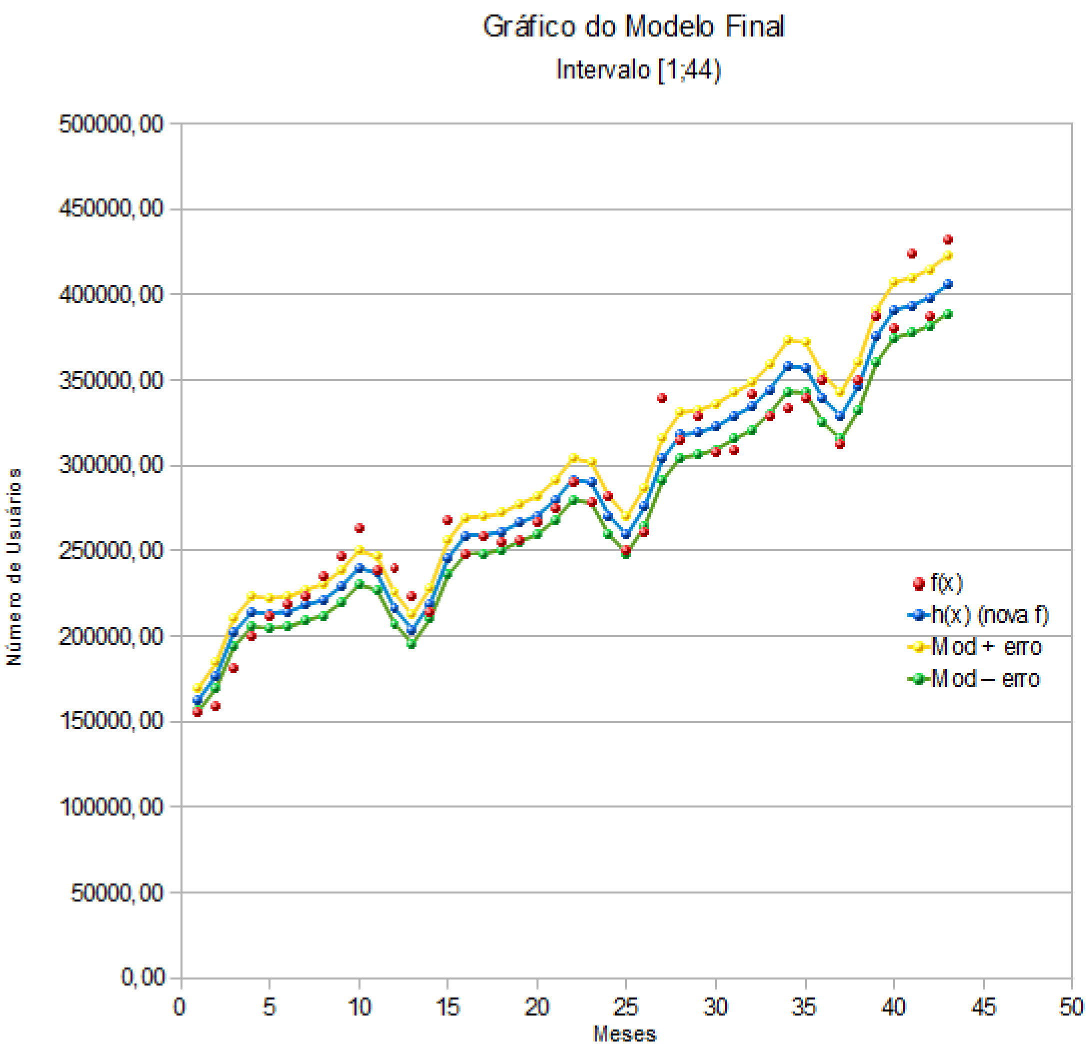

Figura 9 – Gráfico do Modelo Final para o intervalo [1;44)

A Figura 10 mostra o gráfico do Modelo Final para o intervalo $[44;72]$:

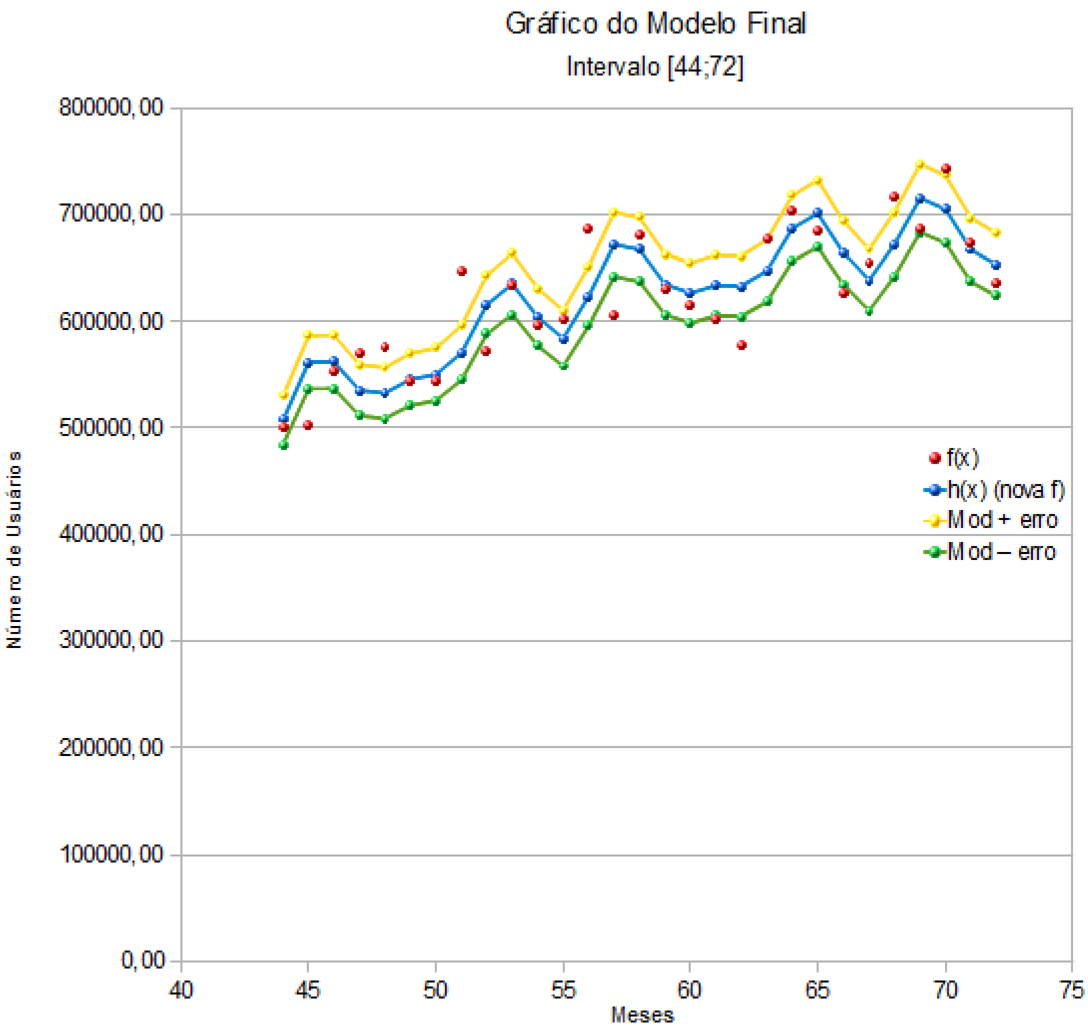

Figura 10 – Gráfico do Modelo Final para o intervalo [44;72]

Note que os dois gráficos apresentam boas aproximações para $f(x)$. As funções crescem e oscilam, e o erro ponderado é o menor entre todos os modelos.

4 Considerações Finais

Neste trabalho, fizemos a construção de um modelo realístico da quantidade de usuários da Linha 9 - Esmeralda, no período compreendido entre janeiro de 2008 e dezembro de 2013. Para Burak (1992) "a modelagem [. . .] constitui-se em um conjunto de procedimentos cujo objetivo é construir um paralelo para tentar explicar, matematicamente, os fenômenos presentes no cotidiano do ser humano, ajudando-o a fazer predições e tomar decisões". Acreditamos que crescimento apontado pelo modelo sugere a necessidade do aumento do sistema metroferroviário e evidencia um impacto causado pela ligação entre as duas linhas na Estação Pinheiros.

A linha observada, até o ano de 2002, só possuía integração com a Linha 8 - Diamante, nas Estações Osasco e Presidente Altino, quando passou a ter integração com a Linha 5 - Lilás, do METRÔ, na Estação Santo Amaro. No período avaliado na pesquisa, as ligações citadas não influenciaram na quantidade de usuários da Linha 9 - Esmeralda.

No entanto, em julho de 2011, houve a integração entre as Linha 4 - Amarela (ViaQuatro) e a Linha 9 - Esmeralda, na Estação Pinheiros, o que impactou a CPTM, notadamente, culminando na divisão do Modelo 7. As interconexões entre as diferentes linhas aumentam o número de pessoas transportadas, tornando o modelo ineficiente e a prova disso foi o fato de necessitarmos dividir o modelo em pré e pós interconexão.

Existe a possibilidade de que outro modelo, sem a necessidade de divisão, se ajuste melhor aos dados obtidos. Todavia, a divisão, dentro dos parâmetros estabelecidos, foi um recurso mais rápido para a definição do Modelo Final, que por esta razão, ficou representativo. Bassanezi (1999) diz que "a validação de um modelo matemático consiste na verificação da aproximação do modelo com a realidade, ou seja, se os dados experimentais ou observados não estão muito longe daqueles fornecidos pelo modelo".

Há, naturalmente, outros métodos para a criação de um modelo. Optamos pelo uso do Método dos Quadrados Mínimos pela facilidade na construção e adaptação. A Série de Fourier foi utilizada por causa do comportamento oscilatório.

De acordo com Bassanezi (2002) "o que se procura numa modelagem é estabelecer um ponto de partida com modelos simples, não comprometedores e que possam ser modificados conforme os objetivos vão sendo ampliados". Esperamos que a pesquisa contribua para a criação de modelos de outras linhas e acreditamos que o passo seguinte é a modelagem da malha metroferroviária paulista completa, com a inclusão de novos parâmetros, inclusive.

A capacidade de usuários por estações pode ser estudada através da modelagem.

Neste caso, outras variáveis precisam ser consideradas, como: capacidade de usuários por trem, tamanho útil da estação em metros quadrados, quantidade de usuários na região.

Também por modelagem, pode ser avaliada a demanda de usuários por estações e as projeções futuras. Isso pode contribuir para que o impacto de integrações de futuras linhas seja menor, ao selecionar estações com baixa demanda. Acreditamos que esta sugestão complementa a anterior.

Sugerimos também, pesquisas sobre o impacto causado pela integração de linhas. Há algumas em construção, como por exemplo, a Linha 6 - Laranja e a Linha 17 - Ouro, e já existem integrações programadas.

Por fim, a ideia pode ser ampliada outros sistemas como o sistema de ônibus, na modelagem de usuários em terminais e corredores de ônibus, sobretudo nas regiões periféricas, em geral, locais em que a malha metroferroviária não está presente.

Referências

ÁVILA, G. *Análise Matemática para Licenciatura*. 3. ed. São Paulo: Edgard Blucher, 2006. Citado na página 57.

BASSANEZI, R. C. Modelagem matemática: uma disciplina emergente nos programas de formaçáo de professores. 1999. Disponível em: <http://www.ime.unicamp.br/~biomat/bio9art_1.pdf>. Citado na página 51.

BASSANEZI, R. C. *Ensino-aprendizagem com modelagem matemática: uma nova estratégia*. São Paulo: Contexto, 2002. Citado na página 51.

BIEMBENGUT, M. S. *Modelagem matemática como método de ensino aprendizagem de matemática em cursos de primeiro e segundo graus*. Dissertação (Mestrado) — Universidade Estadual Paulista, Rio Claro, 1990. Citado na página 23.

BOLDRINI, J. L. et al. *Álgebra Linear*. 3. ed. São Paulo: Harbra, 1980. Citado na página 77.

BRUTON, A. J. *Introdução ao Planejamento dos Transportes*. São Paulo: Universidade de São Paulo, 1979. Citado na página 23.

BURAK, D. *Modelagem Matemática: ações e interações no processo de ensino-aprendizagem*. Tese (Doutorado) — Universidade Estadual de Campinas, Campinas, 1992. Citado na página 51.

FERREIRA, J. Á. T. Cálculo numérico: notas de aula. 2013. Disponível em: <http://www.iceb.ufop.br/decom/prof/bob/com400/Textos_CIC170/Eq%20nao%20linear.pdf>. Citado na página 35.

FIGUEIREDO, D. G. de. *Análise de Fourier e Equações Diferenciais Parciais*. 3. ed. Rio de Janeiro: IMPA, 2014. Citado na página 31.

FRANCO, N. M. B. *Cálculo Numérico*. São Paulo: Pearson, 2007. Citado na página 28.

GERHARDT, T. E.; SILVEIRA, D. *Métodos de Pesquisa*. 2013. Disponível em: <www.ufrgs.br/cursopgdr/derad0005>. Acesso em: 18 abr 2015. Citado na página 24.

GUIDORIZZI, H. L. *Um curso de Cálculo*. 5. ed. Rio de Janeiro: LTC, 2008. v. 2. Citado na página 25.

IBGE. *Estimativas da população dos municípios brasileiros com data de referência em 1 de julho de 2014*. 2014. Disponível em: <http://www.ibge.gov.br/home/presidencia/noticias/pdf/analise_estimativas_2014.pdf>. Acesso em: 14 abr 2016. Citado na página 23.

JUNQUEIRA, F. et al. *Modelagem de trens urbanos em Redes de Petri utilizando o Google Maps*. 2013. Disponível em: <http://www.sbai2013.ufc.br/pdfs/8328.pdf>. Acesso em: 24 abr 2016. Citado na página 24.

MENEZES, E. D. A. G. *Modelagem da demanda e oferta como ferramenta na análise e planejamento de um sistema de transportes*. 2015. Disponível em: <http://www.det.ufc.br/i>. Acesso em: 18 abr 2015. Citado na página 23.

METRO. *Número de usuários transportados no sistema metroferroviário paulista, em 2013, é recorde*. 2014. Disponível em: <http://www.metro.sp.gov.br/noticias/numero-de-usuarios-transportados-nosistema-metroferroviario-paulista-em-2013-e-recorde.fss>. Acesso em: 14 abr 2016. Citado na página 23.

NOBRE, E. A. C.; BIANCHI, M. C. A. P. *Os caminhos e descaminhos do metrô: análise comparativa da implantação da rede metroviária nas metrópoles de São Paulo e México*. 2014. Disponível em: <http://www.fau.usp.br/.../enobre_art4.pdf>. Acesso em: 18 abr 2015. Citado na página 23.

QUINTELLA, M. Visão social e humanística do transporte público sobre trilhos. *Revista Plurale*, p. 20–21, jun-jul 2009. Citado na página 23.

RUGGIERO, M. A. G.; LOPES, V. L. da R. *Cálculo Numérico: aspectos teóricos e computacionais*. 2. ed. São Paulo: Pearson, 1988. Citado na página 32.

SOSSAE, R. C.; ALLEVATO, N. S. G.; RAIMUNDO, S. *A modelagem matemática como estratégia de ensino*. 2013. Disponível em: <http://www.brasilengenharia.com/portal/images/revistas/edicao616/616_beneficios_socioeconomicos.pdf>. Acesso em: 18 abr 2015. Citado na página 24.

SPTRANS. *Indicadores*. 2015. Disponível em: <http://www.sptrans.com.br/indicadores>. Acesso em: 14 abr 2016. Citado na página 23.

VIECILI, C. R. C. *Modelagem matemática: uma proposta para o ensino da matemática*. 2013. Disponível em: <www.repositorio.pucrs.br/dspace/.../1/000380369-Texto%2BCompleto-0.pd>. Acesso em: 18 abr 2015. Citado na página 23.

WALDVOGEL, B. C. et al. *Projeções Demográficas para os Distritos do Município de São Paulo*. 2014. Disponível em: <http://abep.info/files/trabalhos/trabalho_completo/TC-10-42-466-481.pdf>. Acesso em: 14 abr 2016. Citado na página 23.

Anexos

ANEXO A – Demonstração do Teste M de Weierstrass

Esta demonstração do Teste M de Weierstrass pode ser encontrada na obra "Análise Matemática para Licenciatura", de Geraldo Ávila.

É claro que a série de funções converge para uma certa função $|f_n(x)| \leq M_n$, e converge absolutamente, devido à dominação $|f_n(x)| \leq M_n$ e do fato de ser convergente a série $\sum M_n$. A convergência desta série garante que, dado qualquer $\epsilon > 0$, existe N tal que

$$n > N \Rightarrow \sum_{j=n+1}^{\infty} M_j < \epsilon.$$

Então, para todo x em D,

$$n > N \Rightarrow |f(x) - \sum_{j=1}^{n} f_j(x)| = |\sum_{j=n+1}^{\infty} f_j(x)| \leq \sum_{j=n+1}^{\infty} M_j < \epsilon,$$

e isto prova a uniformidade da convergência e conclui a demonstração do teorema.

ANEXO B – Ofício

INSTITUTO FEDERAL DE EDUCAÇÃO,
CIÊNCIA E TECNOLOGIA DE SÃO PAULO
***CAMPUS* SÃO PAULO**

Rua Pedro Vicente, 625, Canindé, São Paulo –SP

São Paulo, 04 de setembro de 2015.

À Companhia Paulista de Trens Metropolitanos

Assunto: **Informações a respeito de pesquisa acadêmica**

Prezado Senhor,

1. Informo a V. Sa. que Karl Willian Sousa Santos, RG nº 36288515-1, é estudante regularmente matriculado no curso de Licenciatura em Matemática deste Instituto.

2. Informo-vos, ainda, que, em cumprimento à determinação curricular de realização de um Trabalho de Pesquisa (elemento obrigatório para a conclusão do curso), o referido licenciando desenvolve seu projeto intitulado "Modelagem do fluxo de usuários da CPTM, em um trecho da linha 9 – Esmeralda, usando redes de Petri", orientado pelo Professor Mestre Lucas Casanova Silva.

Prof. Dr. Henrique Marins de Carvalho
Coordenador do Curso de Licenciatura em Matemática

ANEXO C – Carta

CPTM

Ao

Instituto Federal de Educação, Ciência e Tecnologia de São Paulo
Rua Pedro Vicente, 625 - Canindé - São Paulo - SP

Prezado Prof. Me. Lucas Casanova Silva,

Venho através dessa carta informar que o aluno Karl Willian Sousa Santos, regularmente matriculado no curso de graduação em Matemática, solicitou junto ao Departamento de Planejamento de Transporte – DPTT, da Companhia Paulista de Trens Metropolitanos – CPTM, informações sobre passageiros transportados pela linha 9 – Esmeralda, entre janeiro de 2008 e dezembro de 2013.

Essas informações foram fornecidas em arquivo eletrônico, para fins acadêmicos.

Atenciosamente,

Rodrigo Sartoratto de Alencar
Chefe de Depto. de
Planejamento de Transportes
DPTT

ANEXO D – Tabelas de Dados - 2008

Tabela 4 – 1º semestre de 2008

Estação	Jan	Fev	Mar	Abr	Mai	Jun
Osasco	185364	176619	193658	195654	197728	200660
Pres. Altino	36184	36659	40878	43819	40982	40745
Ceasa	83432	81127	88221	90143	89652	89923
Villa Lobos - Jaguaré	110004	114383	125687	135171	132494	136104
Cidade Universitária	164081	178151	207633	223089	211754	224577
Pinheiros	98485	103063	113831	123435	127900	134630
Hebraica - Rebouças	234709	237354	276122	290748	307292	320206
Cidade Jardim	132172	143384	160269	185759	201892	208769
Vila Olímpia	211608	213634	243272	265977	265133	272231
Berrini	181543	182973	204271	224980	219216	224828
Morumbi	161535	162909	188293	201181	220428	226889
Granja Julieta	109013	108390	128053	138097	143171	150243
Santo Amaro	608271	622485	729755	784673	763443	773522
Socorro	218701	219277	244403	235057	164698	162786
Jurubatuba	201553	208857	233903	236303	199871	196495
Autódromo	70931	73833	87063	95903	72245	74198
Interlagos	—	—	—	29747	109869	119511
Grajaú	—	—	—	93068	352589	382282

Tabela 5 – 2º semestre de 2008

Estação	Jul	Ago	Set	Out	Nov	Dez
Osasco	213055	212896	218325	228434	216025	223823
Pres. Altino	41356	43850	44878	47033	42250	40052
Ceasa	95220	96652	101587	107712	100683	101685
Villa Lobos - Jaguaré	140576	144281	151544	165191	147364	138134
Cidade Universitária	218650	237953	252298	262207	216208	275627
Pinheiros	133938	140807	148127	156510	138272	129951
Hebraica - Rebouças	330870	344526	360483	386108	350868	347791
Cidade Jardim	212710	226049	240212	255789	227559	226394
Vila Olímpia	271513	286235	300214	322986	275872	262234
Berrini	224489	232801	240743	257278	217422	216820
Morumbi	224656	241843	247912	262623	238620	243970
Granja Julieta	154855	156952	167664	180381	153157	151929
Santo Amaro	757741	813569	866346	916962	841052	839194
Socorro	162051	163115	166281	174154	161091	151864
Jurubatuba	200905	207777	212043	231721	213484	212982
Autódromo	79529	83292	96710	95570	100637	87002
Interlagos	130061	141404	150297	164137	153061	153745
Grajaú	429976	452038	485251	535270	507680	520670

ANEXO E – Tabelas de Dados - 2009

Tabela 6 – 1º semestre de 2009

Estação	Jan	Fev	Mar	Abr	Mai	Jun
Osasco	197539	186574	222758	209568	225154	209367
Pres. Altino	38260	37173	46317	44371	46955	45780
Ceasa	96549	92092	104541	100442	102818	98988
Villa Lobos - Jaguaré	138038	132364	158732	142062	152062	152750
Cidade Universitária	250249	251388	335861	274387	264328	253953
Pinheiros	125666	128607	164853	152244	154376	153951
Hebraica - Rebouças	325176	311601	388162	364298	383950	379163
Cidade Jardim	210313	200754	249165	234540	244275	243629
Vila Olímpia	258535	242224	306634	280361	292121	298251
Berrini	213179	196126	245549	226837	235360	238814
Morumbi	219667	208281	263580	244607	255249	253643
Granja Julieta	146570	140152	175299	159408	163776	165661
Santo Amaro	765330	744801	939615	870656	908193	900203
Socorro	146259	134600	165906	153992	159247	157286
Jurubatuba	187098	180087	214350	204063	217013	206980
Autódromo	79307	76322	97863	88373	92162	90589
Interlagos	145445	138866	170854	174698	172408	170818
Grajaú	489389	463790	569942	549041	582710	575008

Tabela 7 – 2º semestre de 2009

Estação	Jul	Ago	Set	Out	Nov	Dez
Osasco	207914	220464	228680	241911	235291	257193
Pres. Altino	44847	46948	49279	50974	48080	48102
Ceasa	98543	100137	99801	107703	102068	101195
Villa Lobos - Jaguaré	151750	157744	164039	166087	161734	150918
Cidade Universitária	224115	249225	247994	258608	245882	262886
Pinheiros	152666	158299	169272	169384	160754	151922
Hebraica - Rebouças	380373	396202	414351	430021	415476	413692
Cidade Jardim	244192	255655	266358	272062	263434	255096
Vila Olímpia	306583	314443	326967	341438	340105	328498
Berrini	244217	251333	259697	264217	249554	236204
Morumbi	259290	266348	272188	283591	271282	274945
Granja Julieta	170720	170629	181254	187445	173039	164493
Santo Amaro	890159	934530	958448	1024884	994044	1039866
Socorro	160560	165692	174458	182571	175969	170803
Jurubatuba	210840	209232	213172	230029	208939	219581
Autódromo	90502	94429	98223	123648	100609	101556
Interlagos	170231	179165	185168	196492	188940	192704
Grajaú	604955	628193	648771	696303	682138	710242

ANEXO F – Tabelas de Dados - 2010

Tabela 8 – 1º semestre de 2010

Estação	Jan	Fev	Mar	Abr	Mai	Jun
Osasco	213363	210692	256550	236847	245523	225606
Pres. Altino	43670	43691	54641	48238	50823	46720
Ceasa	96256	94889	117388	106582	112297	109459
Villa Lobos - Jaguaré	141688	144202	176177	153100	161349	152083
Cidade Universitária	211886	229317	313572	279470	283415	265112
Pinheiros	136487	147846	199891	181345	187505	176880
Hebraica - Rebouças	366477	368994	480116	440501	467814	453950
Cidade Jardim	224328	238745	320840	298139	315909	306453
Vila Olímpia	303641	309313	412086	369138	390779	382807
Berrini	223992	228336	307875	274423	296189	290228
Morumbi	239431	245360	320605	297321	321866	310571
Granja Julieta	151517	154920	203148	188591	204424	203844
Santo Amaro	979884	1060642	1397264	1295517	1305014	1116473
Socorro	153706	158289	197452	183152	186774	181758
Jurubatuba	184807	188255	248635	236633	248705	231193
Autódromo	87066	89778	116817	113112	115527	109500
Interlagos	159611	168212	217300	207910	222435	210455
Grajaú	593116	614714	781396	753919	799027	762865

Tabela 9 – 2º semestre de 2010

Estação	Jul	Ago	Set	Out	Nov	Dez
Osasco	225864	246127	232386	235282	233927	247983
Pres. Altino	45674	50003	48477	49077	48791	47327
Ceasa	111914	117722	115384	117022	116252	118673
Villa Lobos - Jaguaré	151425	165491	159629	164328	168024	167180
Cidade Universitária	257953	307455	288019	284794	291238	281519
Pinheiros	169171	192675	183308	177110	178016	172904
Hebraica - Rebouças	456091	501599	482487	487034	494515	503091
Cidade Jardim	300715	337655	327239	325018	332695	335761
Vila Olímpia	400148	481446	461432	446724	457917	434409
Berrini	287810	316558	297877	294126	293772	301265
Morumbi	307595	342103	335719	339072	340793	360071
Granja Julieta	200418	221390	210954	209866	209592	215490
Santo Amaro	1113388	1222133	1179728	1213037	1221334	1320793
Socorro	182204	198831	192123	194344	193953	195207
Jurubatuba	240985	251809	247032	256395	257042	280189
Autódromo	110350	120932	117564	120823	143500	123278
Interlagos	212541	233499	228766	231025	234907	238538
Grajaú	780379	840005	820981	865111	903969	957198

ANEXO G – Tabelas de Dados - 2011

Tabela 10 – 1º semestre de 2011

Estação	Jan	Fev	Mar	Abr	Mai	Jun
Osasco	209290	214709	233112	224570	241605	212877
Pres. Altino	45268	49861	53151	49493	55094	47129
Ceasa	108808	113013	123692	118382	128746	114598
Villa Lobos - Jaguaré	157248	171230	179788	168278	182965	163789
Cidade Universitária	262880	315814	356456	348824	399262	299551
Pinheiros	157872	185448	199549	191610	234591	385112
Hebraica - Rebouças	458587	513853	573930	580860	635725	526956
Cidade Jardim	298269	344791	375504	367004	408199	393612
Vila Olímpia	379057	432218	466333	451443	519839	497713
Berrini	283102	320308	343414	334420	386767	348687
Morumbi	313868	353005	394157	384836	426748	387711
Granja Julieta	198374	229080	246242	239350	270603	244887
Santo Amaro	1150829	1316369	1499673	1478044	1617326	1468440
Socorro	176556	196321	213454	207243	227726	200576
Jurubatuba	226328	247628	277609	269889	299233	255333
Autódromo	111460	128571	148788	149095	163831	147419
Interlagos	212743	232765	261040	256045	287705	245874
Grajaú	868078	928647	1040382	1031912	1146257	1034715

Tabela 11 – 2º semestre de 2011

Estação	Jul	Ago	Set	Out	Nov	Dez
Osasco	226592	248135	236966	241037	245117	263140
Pres. Altino	50531	60736	58515	58300	58321	57742
Ceasa	122442	130715	127482	130392	129583	130883
Villa Lobos - Jaguaré	181042	200911	203349	210783	211150	203326
Cidade Universitária	186832	222953	199318	197398	196299	179006
Pinheiros	909937	1224012	1444124	2063879	2250216	2371996
Hebraica - Rebouças	479970	515769	475695	453172	440228	422937
Cidade Jardim	372424	426528	400942	403199	399189	379854
Vila Olímpia	516138	596324	575021	621984	621517	558511
Berrini	396067	459467	443746	459307	471775	444282
Morumbi	416093	484320	481311	505789	513665	517030
Granja Julieta	271399	315186	309276	324957	333490	322283
Santo Amaro	1590778	1822158	1818821	1908641	1942625	2011432
Socorro	228826	258427	256064	273367	285079	284032
Jurubatuba	289064	310106	302288	318761	319865	339536
Autódromo	157648	180379	173704	181863	203179	186158
Interlagos	272819	308300	302122	313347	316325	320478
Grajaú	1115353	1235831	1225507	1283017	1318080	1379606

ANEXO H – Tabelas de Dados - 2012

Tabela 12 – 1º semestre de 2012

Estação	Jan	Fev	Mar	Abr	Mai	Jun
Osasco	229970	214936	253811	226545	250168	228728
Pres. Altino	53254	54193	63771	55631	60634	55235
Ceasa	124084	116742	136434	121896	138628	134419
Villa Lobos - Jaguaré	192713	200152	229558	199754	225806	210386
Cidade Universitária	165085	173296	217366	180928	206688	187738
Pinheiros	2338963	2362702	2835425	2538822	2760450	2670189
Hebraica - Rebouças	383642	360143	412688	358717	408038	380736
Cidade Jardim	344169	345228	408405	359048	406035	388858
Vila Olímpia	582927	560640	719393	674803	716819	636480
Berrini	439369	423465	520553	463551	523434	488038
Morumbi	470618	469107	571002	496343	562135	527275
Granja Julieta	317391	303383	363416	317494	357799	332803
Santo Amaro	1830212	1883024	2251967	1991360	2209542	2019973
Socorro	265488	271508	323887	280343	318612	300666
Jurubatuba	273827	282242	330294	286848	324969	295376
Autódromo	169189	177563	209570	180493	201857	188791
Interlagos	294132	296968	345122	295930	332188	313842
Grajaú	1302918	1283576	1443013	1282953	1424658	1375096

Tabela 13 – 2º semestre de 2012

Estação	Jul	Ago	Set	Out	Nov	Dez
Osasco	230754	255518	233642	254901	240817	249090
Pres. Altino	57233	64046	56949	61703	57869	53837
Ceasa	139452	149388	126120	141045	131780	126414
Villa Lobos - Jaguaré	221881	249818	228902	247040	229106	218566
Cidade Universitária	176216	223601	201619	218344	196078	165639
Pinheiros	2704172	3046049	2701865	3013701	2760961	2683424
Hebraica - Rebouças	383201	431768	377701	418864	384492	377681
Cidade Jardim	378843	441756	375415	423600	380296	367349
Vila Olímpia	698434	836357	728341	809821	660683	618590
Berrini	500274	575431	488927	566593	505448	473667
Morumbi	521166	619479	532698	624274	575043	567471
Granja Julieta	341890	396045	333895	380310	345352	328186
Santo Amaro	1962177	2270196	2041141	2306539	2178278	2189566
Socorro	304859	343237	297373	352245	337929	329772
Jurubatuba	292502	337206	304365	338417	339917	338309
Autódromo	187025	216944	194103	217010	224522	194016
Interlagos	314858	353114	312860	348820	325642	320765
Grajaú	1410927	1537909	1384647	1547932	1463143	1475547

ANEXO I – Tabelas de Dados - 2013

Tabela 14 – 1º semestre de 2013

Estação	Jan	Fev	Mar	Abr	Mai	Jun
Osasco	332841	313595	380813	387990	372775	347220
Pres. Altino	80887	77661	90215	94113	90422	82349
Ceasa	130701	121662	139935	140104	140194	122640
Villa Lobos - Jaguaré	227013	222616	249972	260118	257916	221335
Cidade Universitária	162474	160432	195339	207015	198770	176287
Pinheiros	2694054	2619246	3098498	3171216	3100562	2902404
Hebraica - Rebouças	370929	354258	400628	419795	404966	363876
Cidade Jardim	350756	341090	400302	422640	404883	369247
Vila Olímpia	639122	587249	691698	747124	733437	674974
Berrini	502030	465567	553968	603981	579827	539962
Morumbi	541173	530566	641837	671238	650267	589303
Granja Julieta	327860	304160	363017	384604	370470	336866
Santo Amaro	1973712	1925178	2223527	2298969	2205171	2028140
Socorro	302779	295984	348120	365075	355453	319317
Jurubatuba	297098	288762	338201	342717	348039	303466
Autódromo	187863	187583	222813	234915	227156	205108
Interlagos	303215	288963	335398	347955	342170	308939
Grajaú	1412795	1325380	1521953	1568489	1556865	1399033

Tabela 15 – 2º semestre de 2013

Estação	Jul	Ago	Set	Out	Nov	Dez
Osasco	361528	401717	379479	420811	387559	388947
Pres. Altino	84387	95133	92248	101736	90527	83853
Ceasa	136352	144671	138684	151602	139432	133436
Villa Lobos - Jaguaré	246759	268075	271856	282358	258689	237201
Cidade Universitária	176416	218242	201380	212203	189136	165649
Pinheiros	3030157	3300821	3161356	3402102	3079019	2850421
Hebraica - Rebouças	382201	418446	398830	433083	393281	375833
Cidade Jardim	371087	419505	391613	421485	374246	338873
Vila Olímpia	677246	755619	724971	811355	715080	678747
Berrini	554173	613581	587293	644716	554626	496690
Morumbi	600064	670983	643727	698056	630717	591184
Granja Julieta	356836	388235	371449	405037	352256	319564
Santo Amaro	2099862	2253398	2186014	2379566	2132206	2070052
Socorro	341138	372457	355810	382304	343771	323607
Jurubatuba	317556	355188	330571	357527	340123	339906
Autódromo	214104	243116	234501	251012	258155	218440
Interlagos	326948	357177	342815	364558	334140	319874
Grajaú	1516581	1630390	1543430	1666380	1537764	1525876

ANEXO J – Médias

Tabela 16 – Tabela de Médias

Médias	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Janeiro	155977	224032	250607	312145	543219	602072
Fevereiro	159061	214767	260900	349646	543270	578331
Março	181406	267777	340097	388126	646426	677569
Abril	199600	248553	314663	380628	572859	703781
Maio	212242	258453	328632	424012	634914	685519
Junho	218811	255269	307553	387499	596368	627248
Julho	223453	256248	308590	432442	601437	655189
Agosto	234780	266593	341524	500014	685992	717042
Setembro	246717	275451	329395	501903	606698	686446
Outubro	263948	290409	333899	552733	681731	743661
Novembro	238961	278741	340013	569761	629853	672818
Dezembro	240215	282216	350049	576235	615438	636564

ANEXO K – Regra de Cramer

As informações sobre a Regra de Cramer podem ser encontradas na obra "Álgebra Linear", de José Luiz Boldrini, Sueli Costa, Vera Lúcia Figueiredo e Henry Wetzler.

O cálculo da inversa de uma matriz fornece um outro método de resolução de sistemas lineares de equações. Este só se aplica a sistemas lineares em que o número de equações é igual ao número de incógnitas. Suponhamos que desejássemos resolver o sistema linear de n equações e n incógnitas

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ \quad\vdots \qquad \vdots \qquad \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n \end{cases}.$$

Podemos escrever este sistema na forma matricial

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

ou

$$A \cdot X = B$$

onde

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \text{ é a matriz dos coeficientes}$$

e

$$B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \text{ é a matriz dos termos independentes,}$$

e

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ é a matriz das incógnitas.}$$

Para esta equação, suponhamos que $det(A) \neq 0$ e portanto, que A tenha inversa A^{-1}. Então

$$A^{-1}(AX) = A^{-1}B$$

$$(A^{-1}A)X = A^{-1}B$$

$$I_nX = A^{-1}B$$

$$X = A^{-1}B$$

Na forma matricial

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}^{-1} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Usando a relação para matriz inversa, temos

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{det(A)} \begin{bmatrix} \Delta_{11} & \dots & \Delta_{1n} \\ \vdots & & \vdots \\ \Delta_{n1} & \dots & \Delta_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

Então

$$x_1 = \frac{b_1\Delta_1 1 + \ldots + b_n\Delta_n 1}{det(A)}.$$

Mas note que o numerador desta fração é igual ao determinante da matriz que obtemos de A, substituindo a primeira coluna pela matriz dos termos independentes. De fato, usando o desenvolvimento de Laplace, obtemos

$$\begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{vmatrix} = b_1\Delta_{11} + \ldots + b_n\Delta_{n1}.$$

Ou seja

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}}$$

Fazendo as deduções análogas, obtemos

$$x_i = \frac{\begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}}, \text{ para } i = 1, 2, \ldots, n.$$

Observe que no denominador temos o determinante da matriz dos coeficientes ($det(A) \neq 0$), e no numerador aparece o determinante da matriz obtida de A, substituindo a i-ésima coluna pela coluna dos termos independentes. Este método de resolução de um sistema linear de n equações e n incógnitas, que só pode ser aplicado quando o determinante da matriz dos coeficientes for não-nulo, é chamado Regra de Cramer.