Mininimos quadrados

Praciano-Pereira, Tarcisio *

27 de agosto de 2015
préprints da Sobral Matemática
no. 2015.03
Editor Tarcisio Praciano-Pereira
tarcisio@member.ams.org

Resumo

O objetivo deste trabalho é a construção dum exemplo simples do Cálculo Variacional, a determinação da melhor curva que se adapte a uma massa de dados sob a restrição dos quadrados mínimos. Um programa para visualizar o resultado num levantamento de dados é incluído.

Este artigo não está ainda pronto, tem erros! A primeira seção está terminada o trabalho prossegue na correção dos erros das demais seções. São pequenos erros nas contas. Entre as correções uma simplificação sugerida por João Antonio de Sousa ao fazer a simples pergunta, "para que mesmo o vetor $\vec{u}$", que foi então apagado na definição do operador F.

palavras chave: Cálculo Variacional, quadrados mínimos linear, quadrados mínimos não linear

This paper shows a simple example of application of Variational Calculus to solve the problem of fitting curve under the restriction of least squares. A program is provide to show the result on a sample of data.

This paper is work in progress, but the first section is complete. The second and third section will extend the first to reach the non linear case and I got better expressions which made easier the derivatives in both sections 2 and 3. Among the simplifications the vector $\vec{u}$ was erased from the first version because João Antonio de Sousa asked a simple question: "what for?" there was no reason, then $\vec{u}$ was erased from the definition of the operator F.

keywords: linear least square, non linear least square, Variational Calculus

*tarcisio@member.ams.org

1 Quadrados mínimos, o caso linear

O caso “linear”, é também chamado de *quadrados mínimos comum*.

O algoritmo se baseia na existência duma pesquisa em que os dados se encontram coletados numa tabela aqui descrita como, $L = (x_k, y_k)_{k\in\{1,\dots,n\}}$, e se deseja descobrir um função do primeiro grau, $P(x) = ax + b$, tal que

$$y = P(x) = ax + b; \tag{1}$$

$$L = \{(x_k, y_k)_{k\in\{1,\dots,n\}}\}; L \subset \mathbf{R}^{2n} \tag{2}$$

$$F(a,b) = \sum_{k=1}^{n} (P(x_k) - y_k)^2; \sqrt{F(a,b)} = \sqrt{\sum_{k=1}^{n} (P(x_k) - y_k)^2}; \tag{3}$$

$$F(a,b) = \sum_{k=1}^{n} (ax_k + b - y_k)^2 = nb^2 + \sum_{k=1}^{n} (ax_k - y_k)^2 + 2b(ax_k - y_k); \tag{4}$$

em que $\sqrt{F(a,b)}$, equação (eq. 1.3), mede o *desvio padrão* do levantamento relativamente ao modelo P.

Vaiser usada uma hipótese de trabalho que depois será importante e que em nada altera a aplicação algoritmo: a sucessão $(x_k)_{k=1}^n$ é crescente. Na segunda seção isto vai parecer natural quando esta sucessão será dos nós da partição do intervalo $[0, \alpha]; \alpha > 0$. Se a sucessão não for crescente, que seja ordenada paralelamente com a sucessão $(y_k)_{k=1}^n$.

Se os pontos do levantamento satisfizerem à equação duma reta, então

$$y_k = mx_k + l$$

e o valor mínimo de F será alcançado quando $a = m; l = b$, mas esta situação ideal não deve ser esperada e o que se procura é uma reta que melhor represente o *levantamento de dados* no sentido de que o *desvio padrão* seja mínimo. *Desvio padrão* é o nome dado em *Probababilidade e Estatística* à norma euclidiana, $\|\|_2$, que aparece na equação (eq. 1.3).

A equação (eq. 4) define um operador em que a variável, ou elemento do espaço onde o operador está definido, é o polinômio P representando pelos seus coeficientes a, b e o objetivo é encontrar um valor mínimo para este operador no espaço de todos os polinômios do primeiro grau P, é aqui que entra o ponto de vista do *Cálculo Variacional* sob cuja ótica se esta resolvendo este problema.

De imediato se aplicam as condições para obter o mínimo da função F das duas variáveis a, b, definida na equação (eq. 1.3), ou na (eq. 1.4) , o gradiente deve ser zero relativamente às variáveis a, b

$$\frac{\partial F}{\partial a} = 2\sum_{k=1}^{n} x_k\,(ax_k - y_k + b) = 2a\sum_{k=1}^{n} x_k^2 - 2\sum_{k=1}^{n} x_k y_k + 2b\sum_{k=1}^{n} x_k = 0 \tag{5}$$

$$\frac{\partial F}{\partial b} = 2nb + 2\sum_{k=1}^{n} (ax_k - y_k) = 2\left(nb + \sum_{k=1}^{n} ax_k - \sum_{k=1}^{n} y_k\right) = 0; \tag{6}$$

$$\frac{\partial^2 F}{\partial a^2} = 2\sum_{k=1}^{n} x_k^2;\ \frac{\partial^2 F}{\partial b\partial a} = 2\sum_{k=1}^{n} x_k;\ \frac{\partial^2 F}{\partial a\partial b} = 2\sum_{k=1}^{n} x_k;\ \frac{\partial^2 F}{\partial b^2} = 2n; \tag{7}$$

$$H = \begin{pmatrix} \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial b \partial a} \\ \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{pmatrix} = \begin{pmatrix} 2\sum_{k=1}^{n} x_k^2 & 2\sum_{k=1}^{n} x_k \\ 2\sum_{k=1}^{n} x_k & 2n \end{pmatrix} \tag{8}$$

e o determinante do Hessiano deve ser positivo onde o gradiente for zero, caracterizando que se tenha uma forma quadrática positiva definida:

$$H = \det \begin{pmatrix} \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial b \partial a} \\ \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{pmatrix} = \frac{\partial^2 F}{\partial a^2}\frac{\partial^2 F}{\partial b^2} - \frac{\partial^2 F}{\partial b \partial a}\frac{\partial^2 F}{\partial a \partial b}; \tag{9}$$

$$H = 4n\sum_{k=1}^{n} x_k^2 - \left(2\sum_{k=1}^{n} x_k\right)^2 = 4\left(n\sum_{k=1}^{n} x_k^2 - \sum_{k=1}^{n} x_k^2 - 2\sum_{k=1}^{n}\sum_{j=1}^{n} x_k x_j\right); \tag{10}$$

$$H = 4\left((n-2)\sum_{k=1}^{n} x_k^2 + \sum_{k=1}^{n} x_k^2 - 2\sum_{k=1}^{n}\sum_{j=1}^{n} x_k x_j\right); \tag{11}$$

$$H = 4\left((n-2)\sum_{k=1}^{n} x_k^2 + \left(\sum_{k=1}^{n} x_k\right)^2\right) \geq 0; \tag{12}$$

Se houver solução em a, b da equação nula do gradiente, ela será o mínimo requerido, então retornando às equações do gradiente igualado a zero se tem

$$\frac{\partial F}{\partial b} = 2\left(nb + a\sum_{k=1}^{n} x_k - \sum_{k=1}^{n} y_k)\right) = 0 \quad \Rightarrow \quad b = -\frac{a}{n}\sum_{k=1}^{n} x_k + \frac{1}{n}\sum_{k=1}^{n} y_k; \tag{13}$$

$$b = -a\overline{x} + \overline{y}; \tag{14}$$

$$\frac{\partial F}{\partial a} = 2\sum_{k=1}^{n} a x_k^2 - 2\sum_{k=1}^{n} x_k y_k + 2b\sum_{k=1}^{n} x_k = 0 \tag{15}$$

$$a\sum_{k=1}^{n} x_k^2 - \sum_{k=1}^{n} x_k y_k + b\sum_{k=1}^{n} x_k = a\|x\|^2 - <x, y> + b\sum_{k=1}^{n} x_k = 0; \tag{16}$$

$$b = \frac{<x,y> - a\|x\|^2}{\sum_{k=1}^{n} x_k}; \tag{17}$$

$$\frac{<x,y> - a\|x\|^2}{\sum_{k=1}^{n} x_k} = -a\overline{x} + \overline{y}; \tag{18}$$

$$<x, y> - a\|x\|^2 = -a\overline{x}\sum_{k=1}^{n} x_k + \overline{y}\sum_{k=1}^{n} x_k; \tag{19}$$

$$a = \frac{<x,y> - \overline{y}\sum_{k=1}^{n} x_k}{\|x\|^2 - \overline{x}\sum_{k=1}^{n} x_k}; \tag{20}$$

$$P(x) = ax + b; \tag{21}$$

Na equação (eq. 1.17) é preciso que o denominador seja diferente de zero, $\sum_{k=1}^{n} x_k \neq 0$, o que fica resolvido com a translação dos dados do levantamento para o intervalo $[0, \alpha]; \alpha > 0$ que será usada logo no início da próxima seção, como ficou indicado na *hipótese de trabalho* inicial. Se necessário for, se devolvem os dados para o cenário em que foram colhidos, para isto, soma-se uma constante a todos os termos da sucessão $(x_k)_{k=1}^n$ de maneira que $x_1 = 0$.

Na equação (eq. 1.20) é preciso mostrar que denominador não é zero:

$$n\left(\frac{\|x\|^2}{n} - \overline{x}\frac{1}{n}\sum_{k=1}^{n} x_k\right) = n\left(\frac{\|x\|^2}{n} - \overline{x}^2\right) = \|x\|^2 - n\overline{x}^2 = \sum_{k=1}^{n} x_k^2 - n\overline{x}^2; \tag{22}$$

$$\left(\frac{\|x\|^2}{n} - \overline{x}\frac{1}{n}\sum_{k=1}^{n} x_k\right) = \frac{1}{n}\sum_{k=1}^{n} x_k^2 - \overline{x}^2 \geq 0 \tag{23}$$

$$\sum_{k=1}^{n} x_k^2 \geq (\sum_{k=1}^{n} x_k)^2 \tag{24}$$

pela desigualdade de Jensen que aparece na última equação.

Na equação (eq. 21) se tem a expressão do polinômio do primeiro grau que melhor se insere nos dados do levantamenteo L do ponto de vista dos *quadrados mínimos* ou equivalentemente, do *desvio padrão.*

Aqui está um programa escrito em `calc` para obter a função linear que pelo método dos mínimos quadrados melhor se insere num levantamento de dados. A versão que se pode baixar da página [1, MinimosQuadradosLinear.tgz] é a mais recente e revisada. Em seguida vêm as instruções de como fazer uso do programa.

```
mat y[] = {4, 1.5, 6, 8, 10, 7, 14.5, 15, 18, 20}
mat x[] = {-3, -1, 1, 3, 5,  7,  9,   11, 13, 15}
n = size(x);
define escalar(x,y) {
local n = size(x), k = 0, soma = 0;
while(k < n) {
soma = soma + x[k]*y[k];
k++;
}
return(soma);
}
define norma(x) {
local n = size(x), k = 0, soma = 0;
while(k < n) {
soma = soma + x[k]*x[k];
k++;
}
return(sqrt(soma));
}
define media(x) {
local n = size(x), k = 0, soma = 0;
while(k < n) {
                soma = soma + x[k];
                k++;
```

```
        }
return(soma/n);
}
define Soma(x) {
        local n = size(x), k = 0, soma = 0;
        while(k < n) {
                soma = soma + x[k];
                k++;
        }
return(soma);
}

a = (escalar(x,y) - media(y)*Soma(x) )/(power(norma(x),2) - media(x)*Soma(x));
b = (escalar(x,y) - a*power(norma(x),2))/(Soma(x));

define CriaData(x,y) {
        local n = size(x), k = 0, soma = 0;
        local data = fopen("data", "w");
        while(k < n) {
                fprintf(data, "%2.2f %2.2f \n", x[k], y[k]);
                k++;
        }
        fclose(data);
}



define EscreveTransfere() {
local transfere = fopen("transfere", "w");
fprintf(transfere, "P(x) = %2.3f*x + %2.3f \n", re(a),re(b));
fprintf(transfere, "plot \"MinimosQuadradosLinear.data\" with points,\
 P(x), 0 \n");
fprintf(transfere, "pause -2 \"aperte enter para terminar\" \n");
fclose(transfere);
}

define ChamaGnuplot() {
system("gnuplot transfere")
}
define main() {
local a,b,opcao=0;
informa();
        printf( "deseja trocar os valor de x e de y ? \n");
        printf( "Se desejar informe <1> caso contrário informe <0> \n");
        printf( "trocar os valores <0,1> -->"); scanf("%f",opcao);
opcao = sim_nao();
```

```
        if (opcao) {
## trocar os valores de x,y
EntradaMatriz()
}
else {
        EscreveMinimosQuadradosLinear_data(x,y);
        EscreveTransfere(x,y);
        system("joe transfere");
        ChamaGnuplot();
}
}

main();
```

Este programa precisa que no arquivo "`MinimosQuadradosLinear.data`" estejam os dados do levantamento no formato que `gnuplot` possa entender, como os dados distribuidos em duas colunas representando em cada linha, respectivamente, a abcissa e a ordenada dum ponto do plano conforme se pode ver na figura (fig 1), página 5, entretanto, ao rodar o programa com os dados inter-

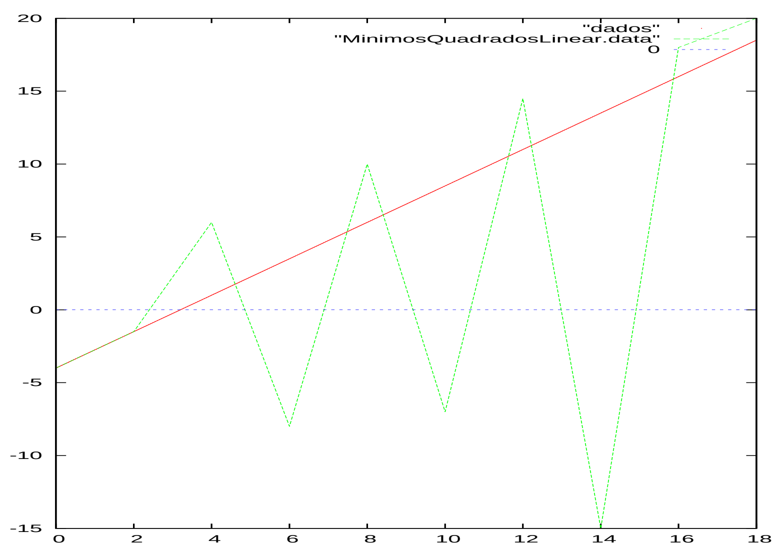

Figura 1: tabela de dados do levantamento

nos do mesmo e os arquivos necessários serão todos criados e lhe servirão de exemplo.

Abra um terminal do `calc` e nele digite:

```
read "MinimosQuadradosLinear.calc"
main();
```

tendo o cuidado de colocar no arquivo `MinimosQuadradosLinear.data` os valores que aparecem na matriz de pontos (x, y) no formato para que `gnuplot` possa entender, confira a figura (fig 1), página 5. Você irá ver a reta que melhor se enquadra nestes dados assim como sua equação estará guardada no arquivo `transfere` que será criado pelo programa.

Você pode ver na figura (fig 2), página 6, o gráfico que corresponde aos dados

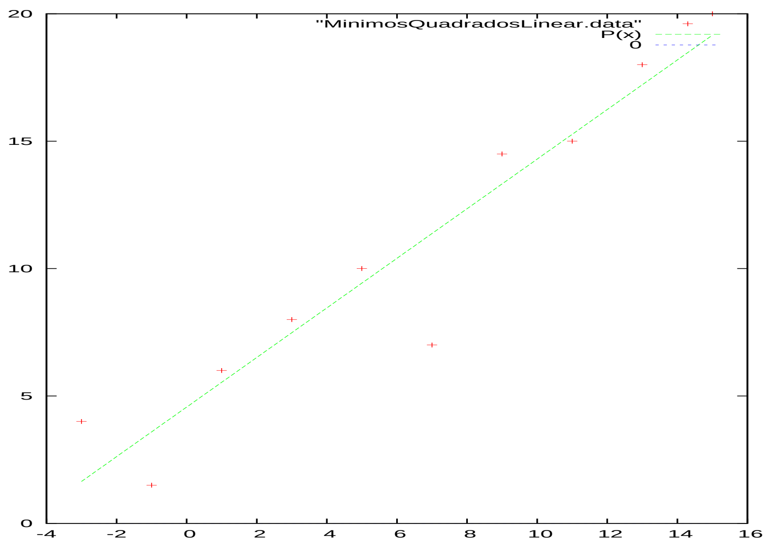

Figura 2: melhor reta para uma massa de dados

tabulados anteriormente.

Você pode baixar os arquivos

- `MinimosQuadradosLinear.tgz`
- `MinimosQuadradosLinear.data`

da página [1] no link `programas`.

Ao rodar o programa `MinimosQuadradosLinear.calc` fique atento á observação que vai aparecer no arquivo contendo os comandos para `gnuplot`, pois será preciso eliminar o simbolo ˜ que `calc` coloca na frente dos números para indicar que o cálculo foi feito aproximadamente. `gnuplot` não irá entender este símbolo e vai apresentar-lhe um erro. O programa também usa o editor `joe`, edite o programa e substitua pelo editor instalado no computador, ou instale `joe`, que é distribuido segundo a `GPL` e é um editor muito avançado.

2 Uma outra forma de calcular

Nesta seção vai ser alterada a metodologia em preparação para a próxima. Você pode saltar para a próxima seção e ao sentir que precisa de um pouco mais de informação retorne para ler esta seção, mas leia rapidamente esta seção para se inteirar das modificações introduzidas relativamente à primeira que são:

1. o intervalo em que o levantamento L se baseia é o intervalo $[0, \alpha]$ resolvendo assim um problema que surgiu na primeira seção, do valor médio da sucessão $(x_k)_k$ que não podia ser zero,

2. a seleção do intervalo $[0,\alpha]$ como *domínio do levantamento*, simplificou muito todos os cálculos uma vez que vou calcular integrais sobre este intervalo, e aqui se aplica a mesma observação feita na primeira seção, esta escolha representa apenas uma translação dos dados que pode ser facilmente revertida, se necessário,

3. se a sucessão for uma sucessão crescente de pontos do intervalo $[0,\alpha]$, com $x_1 = 0; x_n = \alpha$ então se trata da sucessão dos nós duma partição de $[0,\alpha]$ e o seu valor médio será maior do que zero, e se não for, reordene os dados usando $(x_k)_{k=1,\dots,n}$ como referência para a reordenação, quer dizer, se $(x_{s(k)})_{k=1,\dots,n}$ for a sucessão ordenada, atribua $y_{s(k)} := y_k$. Por hipótese então $(x_k)_{k=1,\dots,n}$ é a sucessão dos nós duma partição do intervalo $[0,\alpha]$,

4. uma das contas que se precisa fazer é do cálculo do Hessiano, o determinante da Hessiana, a matriz das derivadas parciais de segunda ordem, onde as entradas são integrais, se o intervalo de integração for $[0,\alpha]$ as contas ficam muito simplificadas, sobretudo as desigualdades, isto sem qualquer perda de generalidade,

5. a função f, do primeiro grau por pedações, portanto um 0-splines tendo por *pontos de precisão* os nós do levantamento L. Em outras palavras, f é uma função do primeiro grau por pedações definida no próximo sistema de equações.

$$(x_k)_{k\in\{1,\dots,n\}} \text{ os nós duma partição de } [0,\alpha]; \tag{25}$$

$$f \text{ é o 0-splines } ; f(x_k) = y_k; x_1 = 0; x_n = \alpha; \tag{26}$$

$$P(x) = ax + b; \tag{27}$$

$$F(a,b) = \int_0^\alpha (P(x) - f(x))^2 dx = \int_0^\alpha (ax + b - f(x))^2 dx; \tag{28}$$

e o objetivo é encontrar os números a, b que definem o polinômio P, exatamente como na primeira seção, que minimiza o operador F definido na equação (2.28).

As derivadas parciais de primeira e segunda ordem são

$$\frac{\partial F}{\partial a} = 2\int_0^\alpha (ax + b - f(x))x dx = 2a\frac{\alpha^3}{3} + b\alpha^2 - 2\int_0^\alpha xf(x)dx; \tag{29}$$

$$\frac{\partial F}{\partial b} = 2\int_0^\alpha (ax + b - f(x))dx = a\alpha^2 + 2b\alpha - 2\int_0^\alpha f(x)dx; \tag{30}$$

$$\frac{\partial^2 F}{\partial a^2} = 2\frac{\alpha^3}{3}; \frac{\partial^2 F}{\partial b \partial a} = \alpha^2; \tag{31}$$

$$\frac{\partial^2 F}{\partial a \partial b} = \alpha^2; \frac{\partial^2 F}{\partial b^2} = 2\alpha; \tag{32}$$

De onde posso deduzir, aplicando a condição do gradiente nulo nos pontos em que se espera ter um extremo,

$$\nabla(F) = \left(\quad 2a\frac{\alpha^3}{3} + b\alpha^2 - 2\int_0^\alpha xf(x)dx; \quad a\alpha^2 + 2b\alpha - 2\int_0^\alpha f(x)dx; \quad \right); \tag{33}$$

$$\nabla(F) = 0 \quad \Rightarrow \quad \begin{cases} 2a\frac{\alpha^3}{3} + b\alpha^2 - 2\int\limits_0^\alpha xf(x)dx = 0; \\ a\alpha^2 + 2b\alpha - 2\int\limits_0^\alpha f(x)dx = 0; \end{cases} \tag{34}$$

$$\begin{pmatrix} 2\frac{\alpha^3}{3} & \alpha^2 \\ \alpha^2 & 2\alpha \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2\int\limits_0^\alpha xf(x)dx \\ 2\int\limits_0^\alpha f(x)dx \end{pmatrix} \tag{35}$$

$$\mathcal{M}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2\int\limits_0^\alpha xf(x)dx \\ 2\int\limits_0^\alpha f(x)dx \end{pmatrix}; det(\mathcal{M}) = \frac{4\alpha^4}{3} - \alpha^4 = \frac{\alpha^4}{3} > 0; \tag{36}$$

$$a = \frac{det(\mathcal{M}_a)}{det(\mathcal{M})}; b = \frac{det(\mathcal{M}_b)}{det(\mathcal{M})}; \tag{37}$$

$$\det(\mathcal{M}_a) = \begin{pmatrix} 2\int\limits_0^\alpha xf(x)dx & \alpha^2 \\ 2\int\limits_0^\alpha f(x)dx & 2\alpha \end{pmatrix}; \det(\mathcal{M}_a) = \int\limits_0^\alpha f(x)(4\alpha x - 2\alpha^2)dx \tag{38}$$

$$\det(\mathcal{M}_a) = 4\alpha I - 2\alpha^2 J; I = \int\limits_0^\alpha xf(x)dx; J = \int\limits_0^\alpha f(x)dx; \tag{39}$$

$$\det(\mathcal{M}_b) = \begin{pmatrix} 2\frac{\alpha^3}{3} & 2\int\limits_0^\alpha xf(x)dx \\ \alpha^2 & 2\int\limits_0^\alpha f(x)dx \end{pmatrix}; \det(\mathcal{M}_b) = \int\limits_0^\alpha f(x)(4\frac{\alpha^3}{3} - 2\alpha^2 x)dx; \tag{40}$$

$$\det(\mathcal{M}_b) = 4\frac{\alpha^3}{3}J - 2\alpha^2 I; \tag{41}$$

$$\begin{cases} a = \frac{12I - 6\alpha J}{\alpha^3}; \\ b = \frac{4\alpha J - 6I}{\alpha^2}; \end{cases} \tag{42}$$

tenho assim os valores dos coeficientes do polinômio em função dos números $I = \int\limits_0^\alpha xf(x)dx, J = \int\limits_0^\alpha f(x)dx, \alpha$.

Passando ao cálculo do Hessiano, que aparece nas equações (2.31) e (2.32) tenho

$$H(F) = \begin{pmatrix} 2\frac{\alpha^3}{3} & \alpha^2 \\ \alpha^2 & 2\alpha \end{pmatrix} = 4\frac{\alpha^4}{3} - \alpha^4 = \frac{\alpha^4}{3} > 0; \tag{43}$$

Então o Hessiano de F é positivo uma vez que não tem sentido que α seja zero, e assim F tem um mínimo quando o gradiente for zero no valores acima calculados para a, b.

Observe que nos cálculos da primeira seção foi necessário impor uma *condição* semelhante: o intervalo $[0, \alpha] \ni x_k$ fosse de números positivos ou que a média das abcissas fosse positiva. É uma pura translação dos dados para tornar possível o cálculo de a, b para depois transladar de volta os dados para recolocá-los no seu *contexto natural*.

2.1 Programa expandido para esta seção

Nesta seção fica mais claro o ganho do uso do Cálculo Variacional, tudo que se precisa é calcular as integrais $I = \int_0^\alpha x f(x) dx$ e $J = \int_0^\alpha f(x) dx$ para obter os coeficientes de P e a forma mais simples de chegar até à fórmula dos coeficientes a, b de P, compare com a primeira seção.

Foi acrescentado ao programa a opção de visualização da reta que melhor que se enquadra nos dados usando a forma de calcular desta seção.

A figura (fig 3), página 9, mostra o resultado da primeira seção, agora cal-

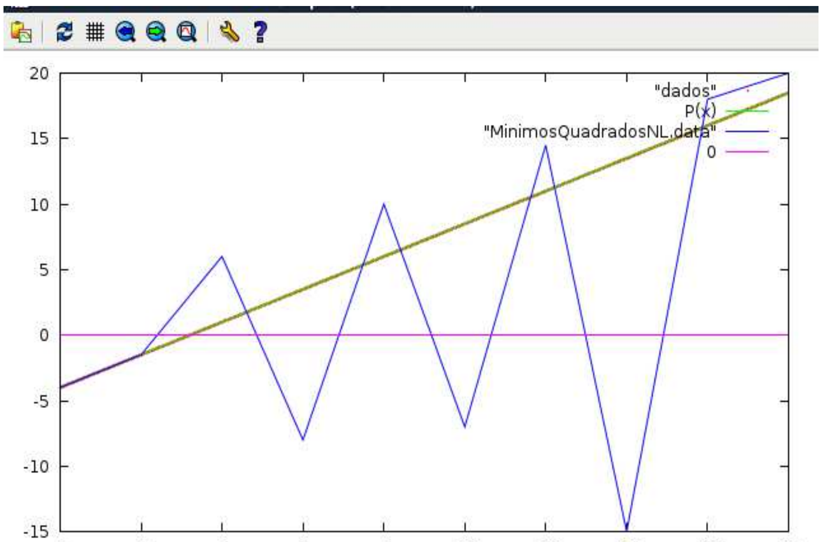

Figura 3: a reta, método Cálculo Variacional

culado o polinôimo $P(x)$ com a metodologia exposta nesta seção.

3 Caso do terceiro grau

Nesta seção será usada a metodologia do Cálculo das Variações para determinar um polinômio do terceiro grau que melhor se adeque a uma massa de dados dum levantamento. A leitora poderá verificar que, com este método, se obtém uma generalização dos casos anteriores e que bastará eliminar os coeficientes de grau maior do que 1, caso eles sejam *despresíveis*, para ter-se a reta que foi discutida nas duas seções anteriores.

Semelhantemente, se o coeficiente do terceiro grau for *despresível*, se pode obter um polinômio do segundo grau que melhor se adeque a uma massa de dados dum levantamento desprezando este coeficiente.

É exatamente isto que caracteriza esta metodologia como inovadora preenchendo o objetivo do artigo. Ao mesmo tempo ficará claro que a metodologia pode facilmente se adaptar para obter um polinômio de grau m qualquer que melhor se adapte a uma massa de dados coletada.

Na última seção se volta à discussão do programa que foi apresentado nas primeiras seções, adaptado para determinar um polinômio de grau m qualquer. Você pode considerar esta seção como o manual do programa já que ela também foi o planejamento do mesmo.

A notação é a mesma da segunda seção cuja leitura não será necessário para esta terceira seção muito além dos primeiros parágrafos, para tomar conhecimento do problema e da notação.

A função f é do primeiro grau por pedações, portanto um 0-splines tendo por *pontos de precisão* os nós do levantamento L e a sucessão $(x_k)_{k\in\{1,\dots,n\}}$, uma sucessão crescente, exatamente os nós da partição do intervalo $[0,\alpha]$. Em outras palavras, f é uma função do primeiro grau por pedações definida no sistema de equações seguintes e o problema consiste em minimizar o funcional F definido na última equação

$$(x_k)_{k\in\{1,\dots,n\}} \text{ os nós duma partição de } [\alpha,\beta]; \tag{44}$$

$$f \text{ é o 0-splines }; f(x_k) = y_k; x_1 = \alpha; x_n = \beta; \tag{45}$$

$$P(x) = ax^3 + bx^2cx + d; \tag{46}$$

$$F(a,b,c,d) = \int_\alpha^\beta (P(x) - f(x))^2 dx = \int_\alpha^\beta (ax^3 + bx^2cx + d - f(x))^2 dx; \tag{47}$$

O objetivo é encontrar os números a, b, c, d que definem o polinômio P, que minimiza o operador F definido na equação (eq. 3.47). exatamente como nas seções anteriores.

Calculando as derivadas parciais,

$$\begin{cases}
\frac{\partial F}{\partial a} = 2\int\limits_{\alpha}^{\beta}(ax^3+bx^2+cx+d-f(x))x^3dx = \\
= 2a\int\limits_{\alpha}^{\beta}x^6dx + 2b\int\limits_{\alpha}^{\beta}x^5dx + 2c\int\limits_{\alpha}^{\beta}x^4dx + 2d\int\limits_{\alpha}^{\beta}x^3dx - 2\int\limits_{\alpha}^{\beta}f(x)x^3dx \\
\frac{\partial F}{\partial b} = 2\int\limits_{\alpha}^{\beta}(ax^3+bx^2cx+d-f(x))x^2dx = \\
= 2a\int\limits_{\alpha}^{\beta}x^5dx + 2b\int\limits_{\alpha}^{\beta}x^4dx + 2c\int\limits_{\alpha}^{\beta}x^3dx + 2d\int\limits_{\alpha}^{\beta}x^2dx - 2\int\limits_{\alpha}^{\beta}f(x)x^2dx; \\
\frac{\partial F}{\partial c} = 2\int\limits_{\alpha}^{\beta}(ax^3+bx^2+cx+d-f(x))xdx = \\
= 2a\int\limits_{\alpha}^{\beta}x^4dx + 2b\int\limits_{\alpha}^{\beta}x^3dx + 2c\int\limits_{\alpha}^{\beta}x^2dx + 2d\int\limits_{\alpha}^{\beta}xdx - 2\int\limits_{\alpha}^{\beta}f(x)xdx; \\
\frac{\partial F}{\partial d} = 2\int\limits_{\alpha}^{\beta}(ax^3+bx^2cx+d-f(x))dx = \\
2a\int\limits_{\alpha}^{\beta}x^3dx + 2b\int\limits_{\alpha}^{\beta}x^2dx + 2c\int\limits_{\alpha}^{\beta}xdx + 2d\int\limits_{\alpha}^{\beta}dx - 2\int\limits_{\alpha}^{\beta}f(x)dx \\
\frac{\partial^2 F}{\partial a^2} = 2\int\limits_{\alpha}^{\beta}x^6dx; \frac{\partial^2 F}{\partial b\partial a} = 2\int\limits_{\alpha}^{\beta}x^5dx = \frac{\partial^2 F}{\partial a\partial b}; \frac{\partial^2 F}{\partial c\partial a} = 2\int\limits_{\alpha}^{\beta}x^4dx = \frac{\partial^2 F}{\partial a\partial c}; \frac{\partial^2 F}{\partial d\partial a} = 2\int\limits_{\alpha}^{\beta}x^3dx = \frac{\partial^2 F}{\partial a\partial d} \\
\frac{\partial^2 F}{\partial b^2} = 2\int\limits_{\alpha}^{\beta}x^4dx; \frac{\partial^2 F}{\partial c\partial b} = 2\int\limits_{\alpha}^{\beta}x^3dx = \frac{\partial^2 F}{\partial b\partial c}; \frac{\partial^2 F}{\partial d\partial b} = 2\int\limits_{\alpha}^{\beta}x^2dx = \frac{\partial^2 F}{\partial b\partial d} \\
\frac{\partial^2 F}{\partial c^2} = 2\int\limits_{\alpha}^{\beta}x^2dx \\
\frac{\partial^2 F}{\partial d\partial c} = 2\int\limits_{\alpha}^{\beta}x^2dx = \frac{\partial^2 F}{\partial c\partial d} \\
\frac{\partial^2 F}{\partial d^2} = 2\int\limits_{\alpha}^{\beta}dx
\end{cases} \tag{48}$$

E quando o gradiente se anula:

$$\left\{\begin{array}{l}
\frac{\partial F}{\partial a}=2\int\limits_{\alpha}^{\beta}(ax^3+bx^2+cx+d-f(x))x^3dx= \\
=2a\int\limits_{\alpha}^{\beta}x^6dx+2b\int\limits_{\alpha}^{\beta}x^5dx+2c\int\limits_{\alpha}^{\beta}x^4dx+2d\int\limits_{\alpha}^{\beta}x^3dx-2\int\limits_{\alpha}^{\beta}f(x)x^3dx \\
\frac{\partial F}{\partial b}=2\int\limits_{\alpha}^{\beta}(ax^3+bx^2cx+d-f(x))x^2dx= \\
=2a\int\limits_{\alpha}^{\beta}x^5dx+2b\int\limits_{\alpha}^{\beta}x^4dx+2c\int\limits_{\alpha}^{\beta}x^3dx+2d\int\limits_{\alpha}^{\beta}x^2dx-2\int\limits_{\alpha}^{\beta}f(x)x^2dx; \\
\frac{\partial F}{\partial c}=2\int\limits_{\alpha}^{\beta}(ax^3+bx^2+cx+d-f(x))xdx= \\
=2a\int\limits_{\alpha}^{\beta}x^4dx+2b\int\limits_{\alpha}^{\beta}x^3dx+2c\int\limits_{\alpha}^{\beta}x^2dx+2d\int\limits_{\alpha}^{\beta}xdx-2\int\limits_{\alpha}^{\beta}f(x)xdx; \\
\frac{\partial F}{\partial d}=2\int\limits_{\alpha}^{\beta}(ax^3+bx^2cx+d-f(x))dx= \\
c=\frac{\int\limits_{\alpha}^{\beta}f(x)dx-b\int\limits_{\alpha}^{\beta}x^2dx-a\int\limits_{\alpha}^{\beta}x^3dx}{\int\limits_{\alpha}^{\beta}xdx}d=\frac{\int\limits_{\alpha}^{\beta}f(x)dx-a\int\limits_{\alpha}^{\beta}x^3dx-b\int\limits_{\alpha}^{\beta}x^2dx-c\int\limits_{\alpha}^{\beta}xdx}{\beta-\alpha}
\end{array}\right. \tag{49}$$

e o Hessiano será

$$H=\begin{pmatrix}
\frac{\partial^2F}{\partial a^2} & \frac{\partial^2F}{\partial b\partial a} & \frac{\partial^2F}{\partial c\partial a} & \frac{\partial^2F}{\partial d\partial a} \\
\frac{\partial^2F}{\partial a\partial b} & \frac{\partial^2F}{\partial b^2} & \frac{\partial^2F}{\partial c\partial b} & \frac{\partial^2F}{\partial d\partial b} \\
\frac{\partial^2F}{\partial a\partial c} & \frac{\partial^2F}{\partial b\partial c} & \frac{\partial^2F}{\partial c^2} & \frac{\partial^2F}{\partial d\partial c} \\
\frac{\partial^2F}{\partial a\partial d} & \frac{\partial^2F}{\partial b\partial d} & \frac{\partial^2F}{\partial c\partial d} & \frac{\partial^2F}{\partial d^2}
\end{pmatrix}=\begin{pmatrix}
2\int\limits_{\alpha}^{\beta}x^6dx & 2\int\limits_{\alpha}^{\beta}x^5dx & 2\int\limits_{\alpha}^{\beta}x^4dx & d\int\limits_{\alpha}^{\beta}x^3dx \\
2\int\limits_{\alpha}^{\beta}x^5dx & 2\int\limits_{\alpha}^{\beta}x^4dx & 2\int\limits_{\alpha}^{\beta}x^3dx & 2\int\limits_{\alpha}^{\beta}x^2dx \\
2\int\limits_{\alpha}^{\beta}x^4dx & 2\int\limits_{\alpha}^{\beta}x^3dx & 2\int\limits_{\alpha}^{\beta}x^2dx & 2\int\limits_{\alpha}^{\beta}xdx \\
2\int\limits_{\alpha}^{\beta}x^3dx & 2\int\limits_{\alpha}^{\beta}x^2dx & 2\int\limits_{\alpha}^{\beta}xdx & 2\int\limits_{\alpha}^{\beta}dx
\end{pmatrix} \tag{50}$$

$$H= \tag{51}$$

$$\mathcal{M}=\begin{pmatrix}
2\frac{\alpha^7}{7} & \frac{\alpha^6}{3} & 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\
\frac{\alpha^6}{3} & 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \\
2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} & \alpha^2 \\
\frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} & \alpha^2 & 2\alpha
\end{pmatrix} \tag{52}$$

$$\det(\mathcal{M})=2\frac{\alpha^7}{7}\det(\mathcal{M}_{11})-\frac{\alpha^6}{3}\det(\mathcal{M}_{21})+2\frac{\alpha^5}{5}\det(\mathcal{M}_{31})-\frac{\alpha^4}{2}\det(\mathcal{M}_{41}); \tag{53}$$

$$\text{para calcular os menores com } \textit{scilab} \tag{54}$$

$$\text{para calcular os menores} \tag{55}$$

$$\text{para calcular } \det(\mathcal{M}_{11}) \tag{56}$$

$$\begin{pmatrix} \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \\ 2\frac{\alpha^3}{3} & \alpha^2 \\ \alpha^2 & 2\alpha \end{pmatrix} \tag{57}$$

$$\det(\mathcal{M}_{11}) = 2\frac{\alpha^5}{5} \begin{vmatrix} 2\frac{\alpha^3}{3} & \alpha^2 \\ \alpha^2 & 2\alpha \end{vmatrix} - \frac{\alpha^4}{2} \begin{vmatrix} \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \\ \alpha^2 & 2\alpha \end{vmatrix} + 2\frac{\alpha^3}{3} \begin{vmatrix} \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \\ 2\frac{\alpha^3}{3} & \alpha^2 \end{vmatrix} \tag{58}$$

$$\det(\mathcal{M}_{11}) = 2\frac{\alpha^5}{5} * \left(2\frac{\alpha^3}{3} * 2\alpha - \alpha^4\right) - \frac{\alpha^4}{2} * \left(\frac{\alpha^4}{2} * 2\alpha - 2\frac{\alpha^3}{3} * \alpha^2\right) + 2\frac{\alpha^3}{3}\left(\frac{\alpha^4}{2} * \alpha^2 - 2\frac{\alpha^3}{3} * 2\frac{\alpha^3}{3}\right) \tag{59}$$

$$\det(\mathcal{M}_{11}) = 2\frac{\alpha^5}{5} * \left(2\frac{\alpha^3}{3} * 2\alpha - \alpha^4\right) - \frac{\alpha^4}{2} * \left(\frac{\alpha^4}{2} * 2\alpha - 2\frac{\alpha^3}{3} * \alpha^2\right) + 2\frac{\alpha^3}{3}\left(\frac{\alpha^4}{2} * \alpha^2 - 2\frac{\alpha^3}{3} * 2\frac{\alpha^3}{3}\right) \tag{60}$$

$$\det(\mathcal{M}_{11}) = 2\frac{\alpha^9}{15} - \frac{\alpha^9}{6} + \frac{\alpha^9}{3*9} = \alpha^9/270) \tag{61}$$

$$\text{para calcular } \det(\mathcal{M}_{21}) \tag{62}$$

$$\begin{pmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ 2\frac{\alpha^3}{3} & \alpha^2 \\ \alpha^2 & 2\alpha \end{pmatrix} \tag{63}$$

$$\det(\mathcal{M}_{21}) = \frac{\alpha^6}{3} \begin{vmatrix} 2\frac{\alpha^3}{3} & \alpha^2 \\ \alpha^2 & 2\alpha \end{vmatrix} - \frac{\alpha^4}{2} \begin{vmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ \alpha^2 & 2\alpha \end{vmatrix} + 2\frac{\alpha^3}{3} \begin{vmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ 2\frac{\alpha^3}{3} & \alpha^2 \end{vmatrix} \tag{64}$$

$$\text{para calcular } \det(\mathcal{M}_{31}) \tag{65}$$

$$\begin{pmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \\ \alpha^2 & 2\alpha \end{pmatrix} \tag{66}$$

$$\det(\mathcal{M}_{31}) = \frac{\alpha^6}{3} \begin{vmatrix} \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \\ \alpha^2 & 2\alpha \end{vmatrix} - 2\frac{\alpha^5}{5} \begin{vmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ \alpha^2 & 2\alpha \end{vmatrix} + 2\frac{\alpha^3}{3} \begin{vmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \end{vmatrix} \tag{67}$$

$$\text{para calcular } \det(\mathcal{M}_{41}) \tag{68}$$

$$\begin{pmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \\ 2\frac{\alpha^3}{3} & \alpha^2 \end{pmatrix} \tag{69}$$

$$\det(\mathcal{M}_{41}) = \frac{\alpha^6}{3} \begin{vmatrix} \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \\ 2\frac{\alpha^3}{3} & \alpha^2 \end{vmatrix} - 2\frac{\alpha^5}{5} \begin{vmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ 2\frac{\alpha^3}{3} & \alpha^2 \end{vmatrix} + \frac{\alpha^4}{2} \begin{vmatrix} 2\frac{\alpha^5}{5} & \frac{\alpha^4}{2} \\ \frac{\alpha^4}{2} & 2\frac{\alpha^3}{3} \end{vmatrix} \tag{70}$$

Para testar usando *scilab* vou usar $\alpha = 2$ e o código para o *scilab* é

```
alpha=2;


M = [2*alpha**7/7, alpha**6/3,   2*alpha**5/5, alpha**4/2;
     alpha**6/3,   2*alpha**5/5, alpha**4/2,   2*alpha**3/3;
     2*alpha**5/5, alpha**4/2,   2*alpha**3/3, alpha**2;
     alpha**4/2,   2*alpha**3/3, alpha**2,     2*alpha]
```

```
det(M)

M11 =  [2*alpha**5/5, alpha**4/2,   2*alpha**3/3;
         alpha**4/2,   2*alpha**3/3, alpha**2;
         2*alpha**3/3, alpha**2,     2*alpha]

M21 = [ alpha**6/3,   2*alpha**5/5, alpha**4/2;
        alpha**4/2,   2*alpha**3/3, alpha**2;
        2*alpha**3/3, alpha**2,     2*alpha]

M31 = [ alpha**6/3,   2*alpha**5/5, alpha**4/2;
        2*alpha**5/5, alpha**4/2,   2*alpha**3/3;
        2*alpha**3/3, alpha**2,     2*alpha]


M41 = [ alpha**6/3,   2*alpha**5/5, alpha**4/2;
        2*alpha**5/5, alpha**4/2,   2*alpha**3/3;
        alpha**4/2,   2*alpha**3/3, alpha**2]


I =(2*alpha**7/7)*det(M11)-(alpha**6/3)*det(M21)+
(2*alpha**5/5)*det(M31)-(alpha**4/2)*det(M41);
```

e o teste foi bem sucedido com $\det(\mathcal{M}) = 0.1733757$ usando o determinante de $\mathcal{M}$, na equação (eq. 3.52) ou do desenvolvimento do determinantes em menores de ordem 2, na equação (eq. 3.70) servindo o teste para verificar a correção das expressões obtidas acima.

Referências

[1] Tarcisio Praciano-Pereira. Programas para cálculo numérico. Technical report, 2009. http://www.calculo-numerico.sobralmatematica.org/programas/.