Escola de Engenharia da Universidade do Minho
Mestrado integrado em Engenharia de Telecomunicações e Informática

TRABALHO PRÁTICO 1
Parte B

Métodos numéricos e Otimização não linear

Bruno Xavier Brás dos Santos, nºa72122

Rui Filipe Ribeiro Freitas, nº a84121

Tiago João Pereira Ferreira, nº a85392

29 de Abril de 2020

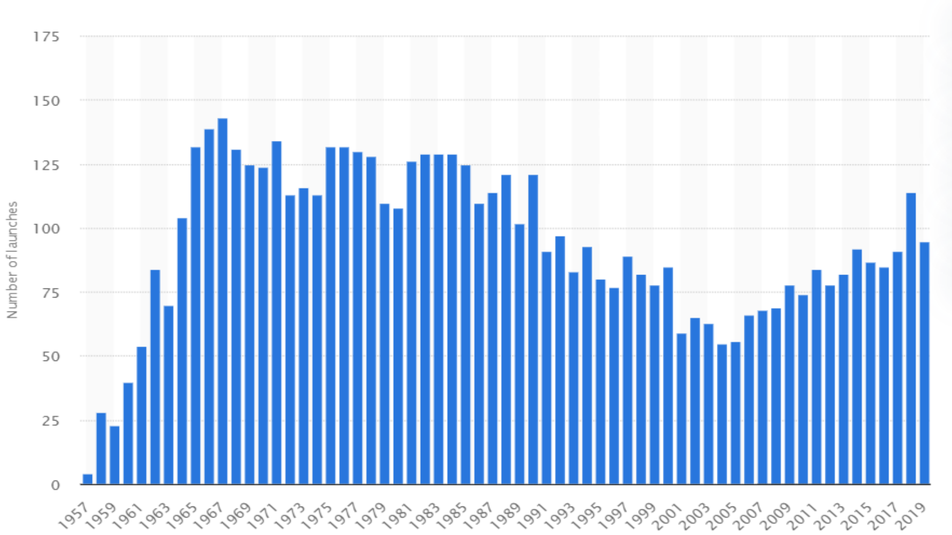
i) Descrição, enquadramento do problema e respetiva formulação matemática

No âmbito da unidade curricular de métodos numéricos e otimização não linear ministrada no curso de MIETI na universidade do Minho foi proposto um projeto que consistia na procura de um determinado problema de ajuste de um modelo a uma função dada por uma tabela matemática, usando a técnica dos mínimos quadrados com auxílio da função *lsqcurvefit* do MATLAB.

Depois de uma análise de vários problemas, o problema escolhido pelo grupo consiste no ajuste por uma função do número de lançamentos de satélites, tripulados e não tripulados, para a órbitra terrestre, por ano, desde o ano de 1957 até ao ano de 2019.

A amostra correspondente, inicialmente analisada em vários textos no site https://www.space.com/, foi posteriormente retirada da tabela encontrada no site http://www.spacelaunchreport.com/logyear.html. Uma vez que a presente tabela contém 63 pontos, e não sendo conveniente apresentar uma tabela destas proporções no presente relatório (limitado a 4 páginas), decidimos representar a amostra através do histograma representado na figura seguinte.

A variável independente, x, corresponde ao ano, cujo domínio vai entre o ano de 1957 e o ano de 2019. A variável dependente, f(x), corresponde ao número de lançamentos efetuados no respetivo ano.



ii) Objetivo e condições de aplicabilidade da função lsqcurvefit

O Objetivo é encontrar um modelo linear nos parâmetros que reflita, na generalidade, o comportamento dos dados usando a técnica dos mínimos quadrados com a ajuda da rotina lsqcurvefit do matlab.

Para aplicar a presente rotina, é necessária uma função que contem o modelo linear cujos parâmetros são as incógnitas a solucionar (modeloLinear), uma aproximação inicial aos parâmetros do modelo [1,1,1,1,1], um vetor com os valores da variável independente (x) e um vetor com os valores da variável dependente (f). No sentido de facilitar a resolução do problema, na variável independente, o ano 1957 corresponderá ao ano 0, o ano de 1958 ao ano 1 e assim sucessivamente. Deste modo, o domínio considerado é [0 a 62] em detrimento de [1957 a 2019]. A sintaxe usada para a rotina lsqcurvit foi:

```
[C,RESNORM,RESIDUAL,EXITFLAG,OUTPUT] = lsqcurvefit('modeloLinear',[1,1,1,1,1],x,f);
```

A rotina devolve vários resultados. No presente projeto apenas guardamos os valores dos parâmetros (C), a raiz quadrada do somatório do quadrado dos resíduos (RESNORM), um vetor com todos os valores dos resíduos para cada ponto da tabela (RESIDUAL), o valor da EXITFLAG e a estrutura OUTPUT.

iii) Ficheiro .m, com a implementação do problema

Na imagem seguinte encontra-se o ficheiro com extensão ‘.m’, onde é possível ver a função colocada no primeiro parâmetro da rotina lsqcurvefit. Nesta função encontra-se o melhor modelo linear que conseguimos obter, ou seja, o modelo que inserido na rotina, esta devolve o menor valor da raiz quadrada do somatório do quadrado do resíduo (RESNORM).

```
function [m] = modeloLinear(c,x)
% modelo linear nos parametros
% C- parametros, x - amostra
m=c(1)*x.^5 + c(2)*x.^2 + c(3)*(x+10.^5) + c(4)*log(1+x.^3) + c(5)*sin(x.^2+1);  % modelo
end
```

Para os inúmeros testes que efetuamos na procura do melhor modelo e dos valores dos parâmetros (C), elaboramos o script que encontra se na figura seguinte.

```
close all; clear; clc;
ficheiro=fopen('satelitesPorAno.txt');
f = fscanf(ficheiro,' %f\n',[1 Inf]);

n=size(f);  % num de elementos de f
x=0:1:n(2)-1;

%construir o grafico original
xaux=x;
yaux1=f;
hold on;
plot(xaux,yaux1,'or');

[C,RESNORM,RESIDUAL,EXITFLAG,OUTPUT] = lsqcurvefit('modeloLinear',[1,1,1,1,1],x,f);

% construir o grafico da aproximacao
yaux2=modeloLinear(C,xaux);
plot(xaux,yaux2,'b');

somatorioQuadradoResiduo=RESNORM^2;

disp('Somatorio do quadrado do residuo: ');  disp(somatorioQuadradoResiduo);
format long;
disp('Valor dos parametros: ');              disp(C);
disp('Valor da EXITFLAG: ');                 disp(EXITFLAG);
disp('Estrutura OUTPUT: ');                  disp(OUTPUT);
```

Na linha 2 do respetivo script encontra-se o nome do ficheiro que contem a tabela com os dados do problema. Na linha 3 é lido para o vetor f, os valores da variável dependente.

iv) Análise de resultados

Através da execução do script mencionado anteriormente e do modelo linear:

$$M = c(1) * \mathrm{x}^5 + c(2) * \mathrm{x}^2 + c(3) * (\mathrm{x} + 10^5) + c(4) * \log(1 + \mathrm{x}^3) + c(5) * \sin(\mathrm{x}^2 + 1)$$

foram obtidos os seguintes resultados:

```
Somatorio do quadrado do residuo:
     4.585738523758171e+07

Valor dos parametros:
   0.000000264211409  -0.093640362396646  -0.000077920870259  19.194303836203360   3.517766030067460

Valor da EXITFLAG:
     3

Estrutura OUTPUT:
    firstorderopt: 4.304872295861132e+02
       iterations: 4
        funcCount: 30
     cgiterations: 0
        algorithm: 'trust-region-reflective'
```

Para este modelo, o somatório do quadrado dos resíduos é de, aproximadamente, 4.5857e+07.

O valor dos 5 parâmetros foram, aproximadamente, 0.264e-6, -0.9364, 0.7792e-4, 19.1943 e 3.5177, respetivamente para c(1), c(2), c(3), c(4) e c(5).

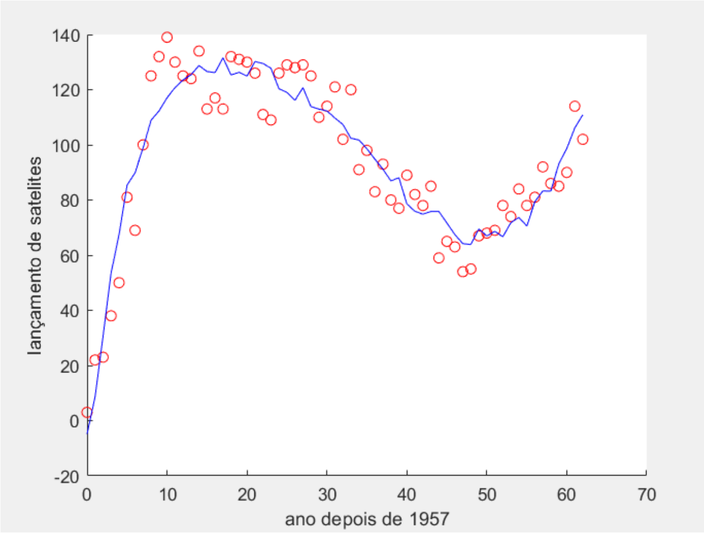
Através do valor da EXITFLAG concluímos que o algoritmo convergiu, mas reconhecendo que o valor ideal de EXITFLAG seria 1. O número de interações foram 4.

De inúmeros modelos testados, este foi o que teve um menor valor do somatório do quadrado dos resíduos, portanto, o que melhor se ajustou ao número de satélites lançados por ano desde o ano 1957 até ao ano de 2019. O modelo, já com os parâmetros obtidos, é o seguinte:

$$M1 = 0.264 * 10^{-6} * \mathrm{x}^5 - 0.9364 * \mathrm{x}^2 + 0.7792 * 10^{-4} * (\mathrm{x} + 10^5) + 19.1943 * \log(1 + \mathrm{x}^3) + 3.5177 * \sin(\mathrm{x}^2 + 1)$$

Na figura seguinte podemos ver, a vermelho os dados retirados da tabela da amostra anunciada no ponto i), e a azul o modelo linear M1 cujo objetivo é ajustar-se aos dados representados na tabela.

Analisando o gráfico, nos primeiros anos os números de satélites aumentaram significativamente, o que é fácil de explicar, uma vez por volta dos anos 1966 estávamos em plena guerra fria e alguns países debatiam-se por uma corrida espacial intensa. Seguidamente o número anual de lançamentos sofreu um decréscimo e, nos últimos anos, o número anual de lançamentos têm aumentado continuamente.



Os modelos M2 e M3 seguintes correspondem a outros dois modelos lineares que permitem ajustar o comportamento dos dados das tabelas.

$$M2 = c(1) * \mathrm{x}^4 + c(2) * \mathrm{x}^2 + c(3) * \mathrm{x} + c(4) * (\mathrm{x} + 10^5) + c(5) * \log(1 + \mathrm{x}^3)$$

$$M3 = c(1) * \mathrm{x}^5 + c(2) * \mathrm{x} + c(3) * \mathrm{x}^2 + c(4) * \cos(\mathrm{x}^2 + 1)$$

O modelo M2 mantem o mesmo número de parâmetros que o modelo M1, mas é substituída a função sinusoidal por uma função polinomial de grau 1. Na figura seguinte, e numa tentativa de representar os 3 modelos num mesmo gráfico (M2 representado a vermelho), podemos ver que o comportamento deste modelo é mais linear do que o do modelo M1. Para este modelo, o valor do somatório do quadrado do resíduo foi de 4.7845e+07, muito próximo do modelo M1, e os valores dos parâmetros foram c(1)=0.272e-4, c(2)=0.1748, c(3)=3.2303, c(4)=0.4278e-4 e c(5)=15.0340 .

O modelo M3 apresenta apenas 4 parâmetros e, dos 3 modelos, é o que apresenta um valor do somatório do quadrado do resíduo superior: 2.8613e+08 (M3 representado a verde).