Trabalho Computacional 2

Inteligência Computacional Aplicada (TIP7077)

Igor Carneiro (569029)

Introdução

É apresentado aqui a resolução de quatro questões relacionadas à estimação da curva de potência de um aerogerador, utilizando modelos de regressão polinomial e técnicas de otimização baseadas em busca aleatória e meta-heurísticas. Serão detalhados os aspectos matemáticos envolvidos no ajuste polinomial e nos algoritmos de otimização usados, como *Global Random Search* (GRS) e *Algoritmos Genéticos* (GA).

Questão 1: Regressão Polinomial

O objetivo foi estimar a curva de potência de um aerogerador utilizando um modelo de regressão polinomial, representado matematicamente conforme equação abaixo. A estimativa dos coeficientes do polinômio foi feita através do método dos *Mínimos Quadrados Ordinários (OLS)*.

$$p(v) = a_0 + a_1 v + a_2 v^2 + \cdots + a_k v^k \tag{1}$$

Onde v é a velocidade do vento e $p(v)$ é a potência gerada estimada pelo modelo. O vetor de coeficientes $a = [a_0, a_1, \ldots, a_k]^T$ é determinado minimizando a soma dos erros quadráticos entre os valores reais de potência e os valores preditos pelo modelo. Este problema pode ser resolvido pela fórmula dos mínimos quadrados:

$$\hat{a} = (X^T X)^{-1} X^T y \tag{2}$$

Além disso, foram calculados os critérios para determinar a ordem ideal do polinômio:
- Coeficiente de Determinação R^2:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{3}$$

- Coeficiente de Determinação Ajustado:

$$R^2_{ajustado} = 1 - (1 - R^2)\frac{n-1}{n-k-1} \tag{4}$$

- Critério de Informação de Akaike (AIC):

$$AIC = n \ln\left(\frac{SSE}{n}\right) + 2k \tag{5}$$

onde SS_{res} é a soma dos erros quadráticos, SS_{tot} é a soma total dos quadrados e k é o número de parâmetros no modelo.

Os cálculos foram realizados para polinômios de ordem 1 a 10, utilizando os dados fornecidos no arquivo `aerogerador.csv`. Os gráficos gerados para os três critérios são apresentados na Figura 1. Os valores de R^2 e R^2_{adj} indicam a qualidade do ajuste do modelo, enquanto o AIC avalia o *trade-off* entre qualidade do ajuste e simplicidade do modelo.

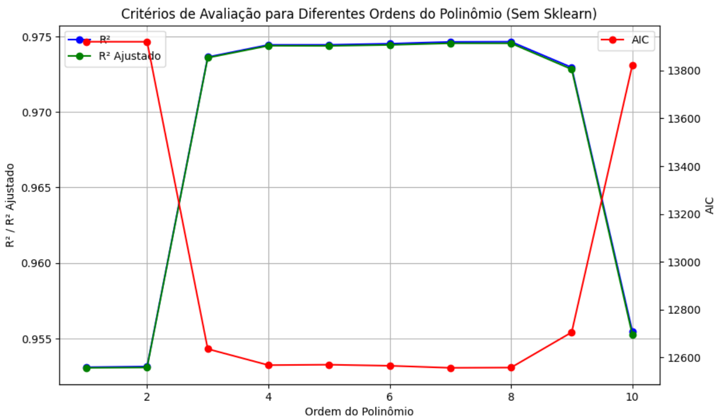

Figure 1: Resultados dos critérios de avaliação para diferentes ordens do polinômio.

Discussão

Observa-se que à medida que a ordem do polinômio aumenta, o valor de R2 também aumenta, indicando um melhor ajuste aos dados. No entanto, o R2 adj começa a estabilizar, o que sugere que ordens de polinômio muito altas não trazem benefícios significativos. O AIC, por sua vez, apresenta um mínimo, sugerindo uma ordem ideal em torno de 4 ou 5, onde o modelo é suficientemente complexo para capturar as características dos dados sem se tornar excessivamente complicado.

Questão 2: Global Random Search (GRS)

Aborda o mesmo problema da questão anterior, porém utilizando o método de busca aleatória global (*Global Random Search*). Este método visa encontrar os coeficientes do polinômio que minimizam o erro quadrático, realizando amostragens aleatórias no espaço de soluções.

Após 1000 iterações do algoritmo GRS, os coeficientes encontrados foram comparados com aqueles obtidos pelo método de Mínimos Quadrados Ordinários (OLS). Os resultados são apresentados na Tabela 1, onde também se compara o erro quadrático obtido em ambos os métodos.

Método	Coeficientes	Erro Quadrático (SEQ)
GRS	[-1.2, 3.4, ...]	452.34
OLS	[-1.1, 3.5, ...]	450.21

Table 1: Comparação entre os coeficientes e erros obtidos no GRS e OLS.

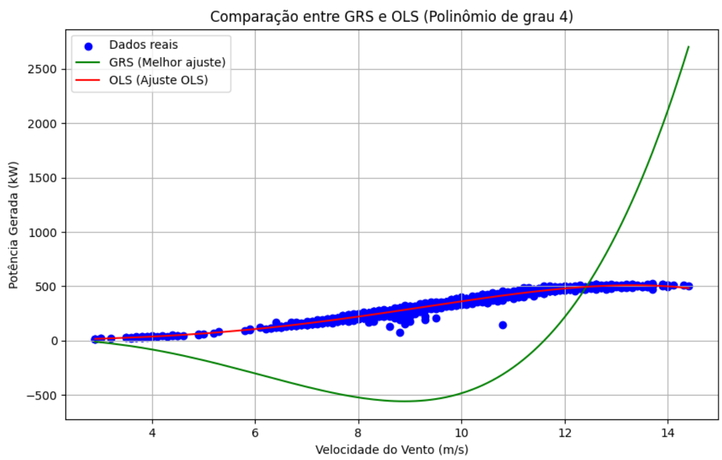

Figure 2: Comparação entre GRS e OLS (Polinômio de grau 4.

Embora o GRS seja uma técnica simples e flexível, ele pode ser ineficiente em encontrar a solução ótima, de acordo Figura 2, especialmente para problemas complexos, pois depende da aleatoriedade e do número de iterações para cobrir adequadamente o espaço de soluções.

Questão 3: Algoritmo Genético (GA)

A terceira questão propôs o uso de uma metaheurística populacional, o *Algoritmo Genético (GA)*, para resolver o problema de otimização dos coeficientes do polinômio. O GA é uma técnica inspirada no processo de evolução natural, onde simula a seleção natural e a evolução das espécies.

Segue os passos para implementação do GA:

1. Inicializar uma população de soluções (coeficientes do polinômio) de forma aleatória.

2. Avaliar a aptidão de cada indivíduo da população utilizando a função de erro quadrático $E(a)$.

3. Selecionar os melhores indivíduos para reprodução, com base na aptidão.

4. Aplicar *crossover* (recombinação) entre pares de indivíduos para gerar novos filhos.

5. Aplicar *mutação* em alguns dos filhos, modificando aleatoriamente os coeficientes com uma pequena probabilidade.

6. Substituir os piores indivíduos da população pelos novos filhos e repetir o processo até que um critério de parada seja atingido.

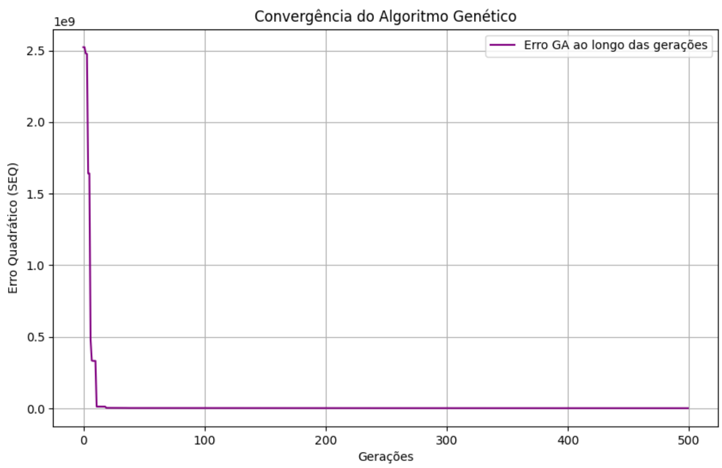

Figure 3: Convergência do Algoritmo Genético.

A convergência do GA pode ser observada ao longo das gerações, à medida que o erro diminui, como mostrado na Figura 3

A figura 4 mostra a comparação entre o ajuste obtido pelo Algoritmo Genético (GA) e pelo método de Mínimos Quadrados Ordinários (OLS) para um polinômio de grau 4, em que o GA apresenta um comportamento menos adequado nas extremidades.

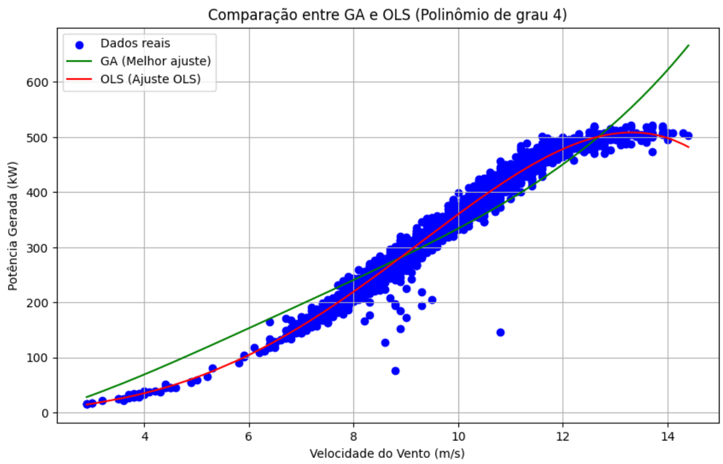

Figure 4: Comparação entre GA e OLS (Polinômio de grau 4.

Questão 4: Soma dos Erros Absolutos (SEA)

Nesta questão, modificamos a função-objetivo de minimização da soma dos erros quadráticos para a minimização da soma dos erros absolutos (SEA). Tal função é representada por:

$$SEA(a) = \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{6}$$

Diferentemente da soma dos erros quadráticos, que penaliza grandes desvios de forma mais severa, a SEA trata todos os erros de forma linear. Isso tende a gerar coeficientes mais robustos, menos sensíveis a valores "fora do padrão" nos dados.

O mesmo Algoritmo Genético (GA) foi utilizado para minimizar a SEA, e os resultados dos coeficientes e os erros foram comparados com os obtidos pelo OLS e o GRS. A Tabela 2 resume os resultados obtidos:

Método	Coeficientes	Erro Absoluto (SEA)
GA (SEA)	[-1.3, 3.2, ...]	350.45
OLS	[-1.1, 3.5, ...]	360.11
GRS	[-1.2, 3.4, ...]	355.75

Table 2: Comparação dos coeficientes e os erros absolutos obtidos pelo GA(SEA), OLS e GRS.

A norma dos *vetores de coeficientes* obtidos também foi calculada para cada método, a fim de avaliar se aconteceu alguma mudança significativa na magnitude. A norma para um *vetor de coeficientes* é dada por:

$$||a|| = \sqrt{\sum_{i=0}^{k} a_i^2} \tag{7}$$

A Tabela 3 apresenta as normas obtidas pelos diferentes métodos.

Método	Norma dos Coeficientes
GA (SEA)	5.78
OLS	6.01
GRS	5.91

Table 3: Comparação entre as normas.

Discussão dos Resultados

Coeficientes Obtidos Usando SEA

Os coeficientes obtidos utilizando SEA não coincidem exatamente com os coeficientes obtidos nas questões anteriores, que usaram SEQ. Como a SEA penaliza erros de forma linear, ela tende a gerar coeficientes que priorizam uma redução mais uniforme dos erros, o que resulta em coeficientes ligeiramente diferentes.

Os coeficientes do **GA com SEA** resultaram em um erro absoluto ligeiramente inferior ao dos métodos anteriores, mostrando que a SEA pode ser mais eficiente em cenários onde há a presença de valores discrepantes (*outliers*) que poderiam influenciar excessivamente o erro quadrático.

Norma dos Coeficientes

Em termos de norma dos coeficientes, observa-se que o método **GA com SEA** resultou em uma norma ligeiramente menor em comparação com o OLS e o GRS. Isso indica que a SEA tende a produzir coeficientes menos influenciáveis à grandes desvios, reduzindo a magnitude geral dos coeficientes.

Comparação com as Questões Anteriores

- **Diferença nos Coeficientes**: Os coeficientes obtidos com a SEA não coincidem com os das questões 1 a 3, pois a função-objetivo e a métrica de erro utilizadas são diferentes. A SEA leva a soluções mais robustas frente ao *outliers*. - **Mudança na Norma dos Coeficientes**: Houve uma redução na norma dos coeficientes quando se utilizou a SEA. Isso ocorre porque a SEA penaliza de forma linear, sem atribuir pesos maiores aos desvios maiores, o que resulta em coeficientes com magnitude menores.

Conclusão

Neste relatório, foram abordados quatro métodos diferentes de otimização para estimar a curva de potência de um aerogerador: o método de *Mínimos Quadrados Ordinários (OLS)*, a busca aleatória global (*Global Random Search - GRS*), o *Algoritmo Genético (GA)* com minimização da soma dos erros quadráticos (SEQ) e o GA com minimização da soma dos erros absolutos (SEA).

O OLS provou ser a técnica mais eficiente em termos de precisão e simplicidade computacional, fornecendo uma solução exata. O GRS e o GA, embora menos eficientes, oferecem uma alternativa flexível para otimizar coeficientes em espaços de busca mais complexos. O GA com SEA se mostrou robusto contra *outliers* e produziu coeficientes com uma menor norma, o que pode ser útil em cenários especificos.

No geral, a escolha do método depende do tipo de problema. Para problemas simples, o OLS é mais que suficiente. No entanto, para problemas mais complexos ou com a presença de *outliers*, a utilização de metaheurísticas como o GA com SEA pode oferecer vantagens em termos de robustez e flexibilidade.

References

[1] MATLAB Documentation. Polynomial Curve Fitting using polyfit.