

2COP507 - Tópicos em Reconhecimento de Padrão

Prof. Sylvio Barbon Jr

Abril 2020

Atividade II - *Ajuste de Curva Polinomial*

Teórica:

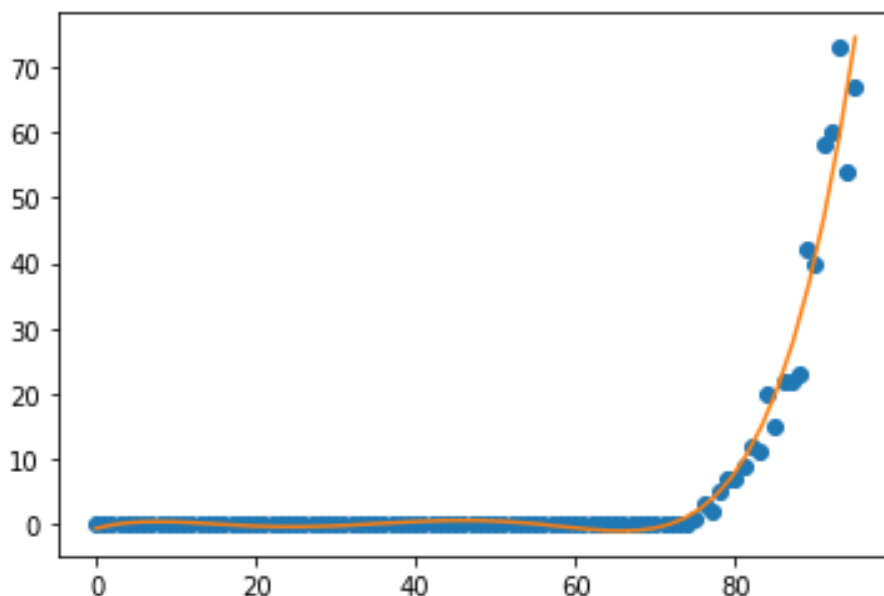
Uma função polinomial é um dos grandes aliados para modelar um comportamento ou padrão. Variando seus parâmetros é possível estimar comportamentos e até prevê-los, como por exemplo o que estamos fazendo atualmente para avaliar o impacto do Coronavírus na estrutura de saúde do Brasil. Par compreender como funciona a função polinomial e seu ajuste, estude a primeira seção da Referência 1 para discutir os seguintes tópicos:

- Explique os componentes da Equação 1.1 e como ela poderia ser apresentada.
- Explique a Equação 1.2 e qual a relação dela com a 1.1.
- O que é Over-fitting?
- Qual a importância de M para evitar o Over-fitting?
- Qual a relação entre:
 - Um bom modelo preditivo (função bem ajustada);
 - Muitas amostras;
 - Over-fitting;
 - Poucas amostras;
 - Under-fitting;
 - Processo de regularização;

Prática:

O exercício prático deste tema é o ajuste de uma função polinomial para modelar as mortes relacionadas ao COVID-19 no Brasil. Para realizar este exercício, sugiro o uso da linguagem Python, com a função *polyfit* do *Numpy*. A base de dados foi encaminhada junto com o descritivo da atividade. O objetivo da atividade é encontrar qual o grau do polinômio que melhor se ajusta aos dados atuais. Sugestões:

1. Abra a base de dados e extraia apenas os dados de morte (coluna "deaths");
2. Plote os dados e verifique se parecem com a figura abaixo (pontos azuis);
3. Separe o conjunto de dados em índices pares (treino) e ímpares (teste);
4. Construa uma função de ajuste usando o *polyfit* com os dados de treino;
5. Crie um método para verificar o Erro Médio Quadrático entre a função encontrada e os dados de teste.
6. O menor erro encontrado é a solução.



Avaliação:

O critério de avaliação será completude das descrições para a atividade teórica. Qualidade das discussões e quantidade de acertos para a atividade prática.

Referência:

- [1] Bishop, Christopher M. Pattern recognition and machine learning. springer, 2006.
- [2] Witten, Ian H., and Eibe Frank. "Data mining: practical machine learning tools and techniques with Java implementations." *Acm Sigmod Record* 31.1 (2002): 76-77.