



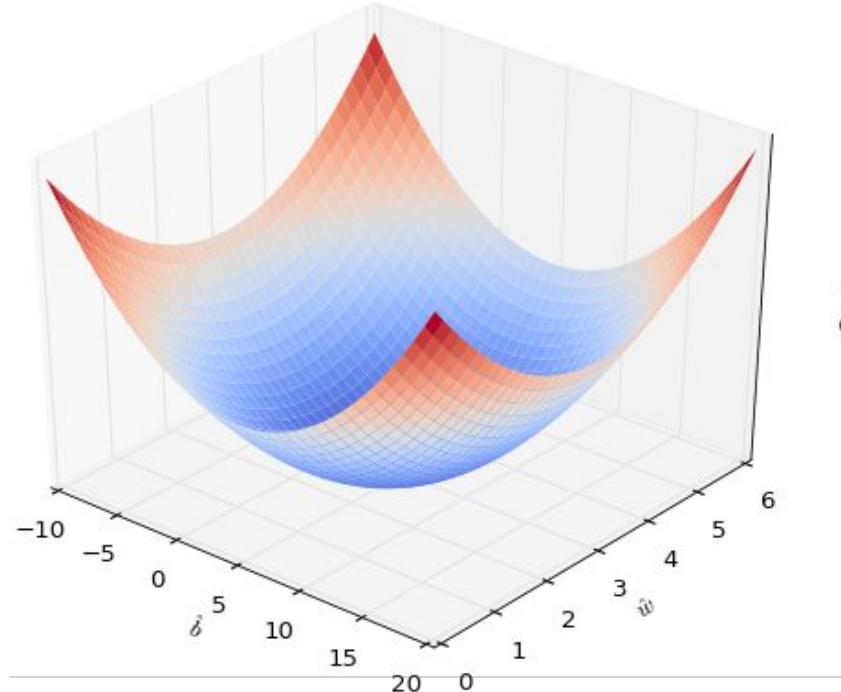
Gradiente Descendente

Objetivos

- Conhecer o algoritmo de Gradiente Descendente
- Entender as vantagens e desvantagens dessa abordagem para o problema de regressão linear
- Introduzir outras soluções para o problema de otimização

Gradiente Descendente (GD)

- Algoritmo de otimização muito genérico capaz de encontrar ótimas soluções para uma ampla gama de problemas
- GD ajusta iterativamente os parâmetros para minimizar uma função de custo



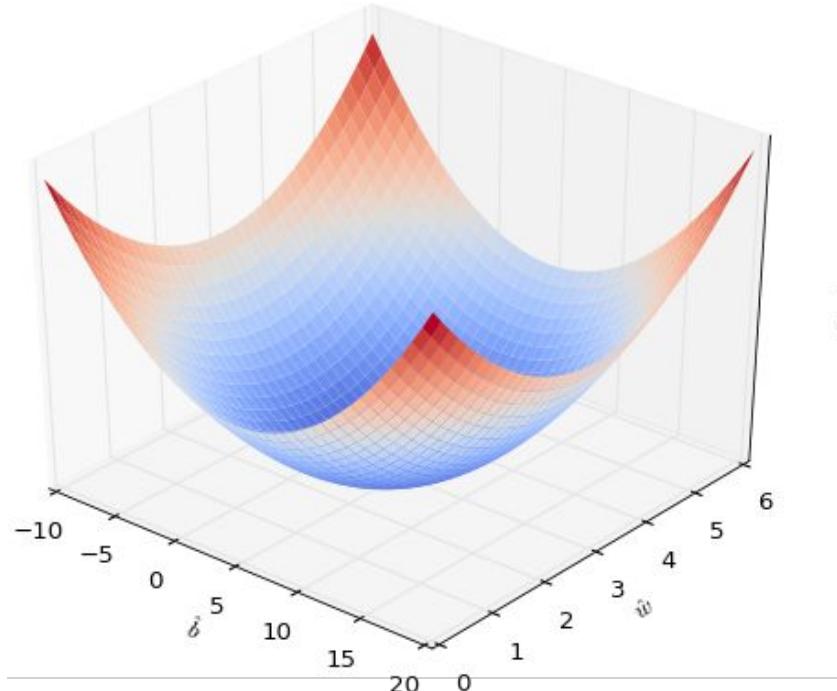
Gradiente Descendente (GD)

Passos:

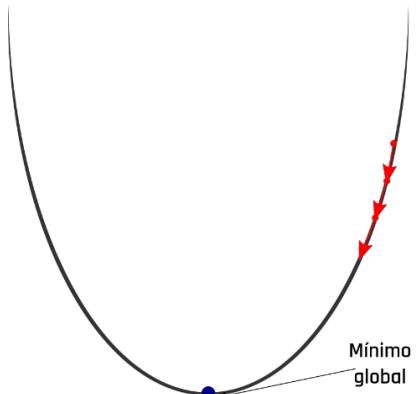
1. Inicia Θ com pesos aleatórios
2. Mede o gradiente local da função de erro em relação ao vetor de parâmetro Θ (derivada parcial)
3. Quando o gradiente for zero, chegamos em um mínimo

Importante:

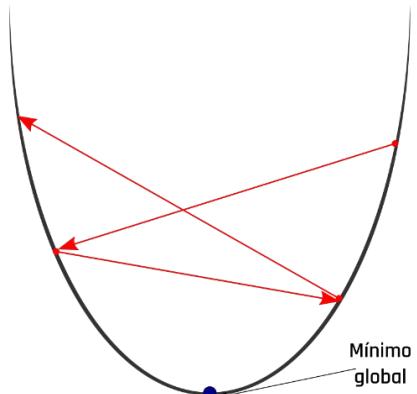
- O tamanho do passo (taxa de aprendizado) se for pequeno demora muito para o algoritmo convergir
- Se for muito grande, não encontramos uma boa solução



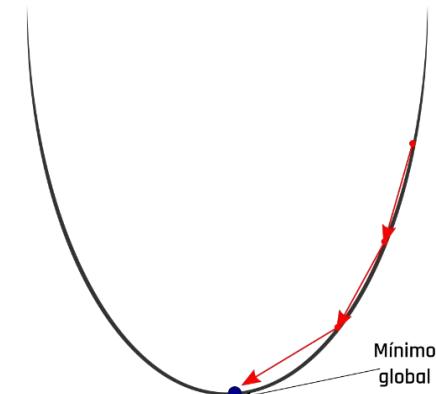
Taxa de aprendizado



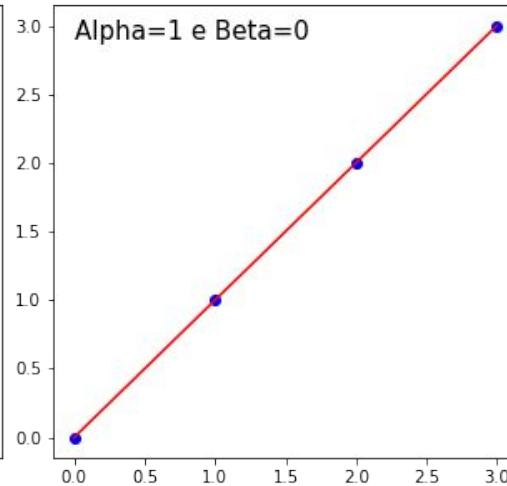
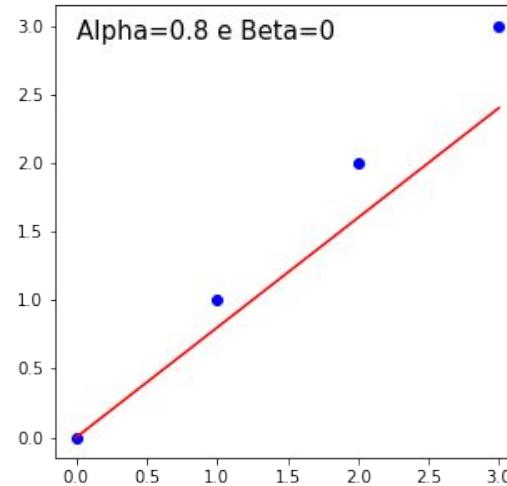
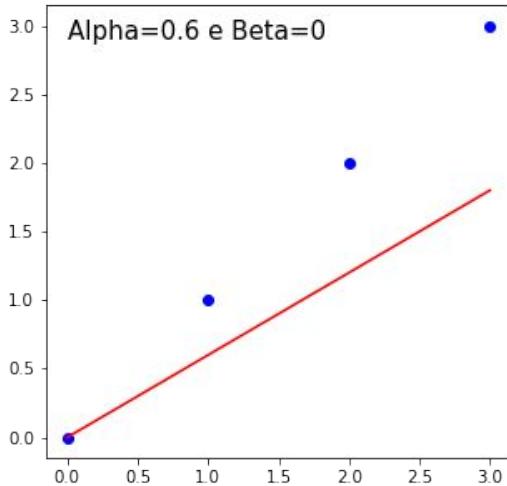
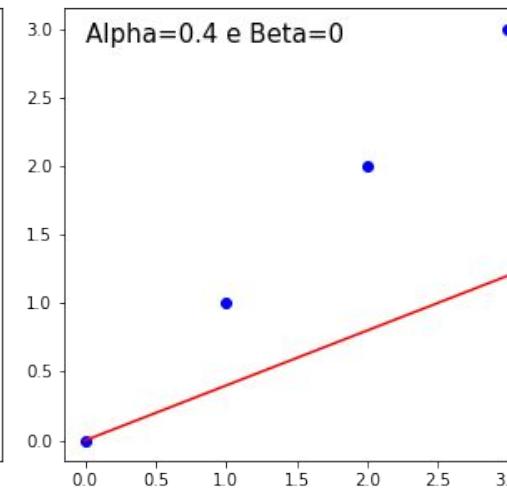
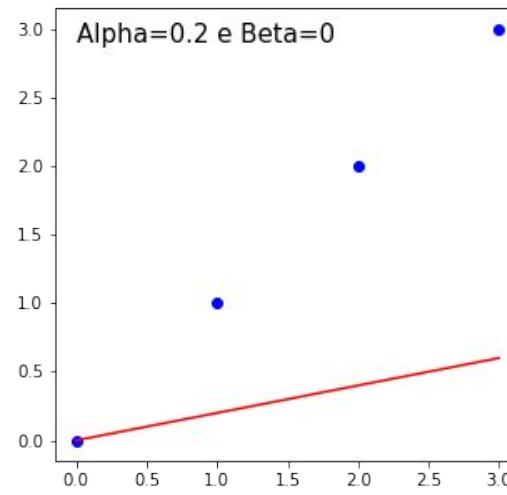
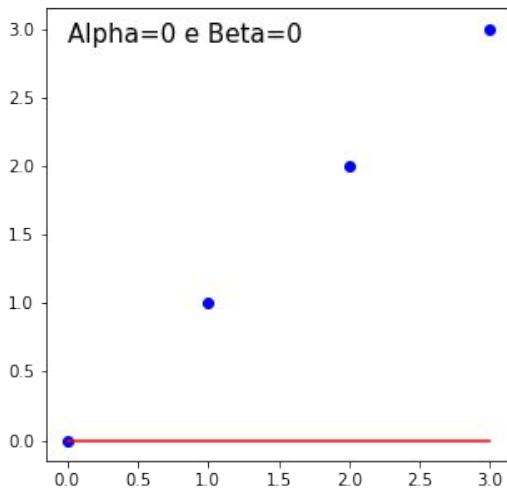
Taxa de aprendizado pequena



Taxa de aprendizado grande



Taxa de aprendizado ideal



Gradiente Descendente (GD)

Vantagens

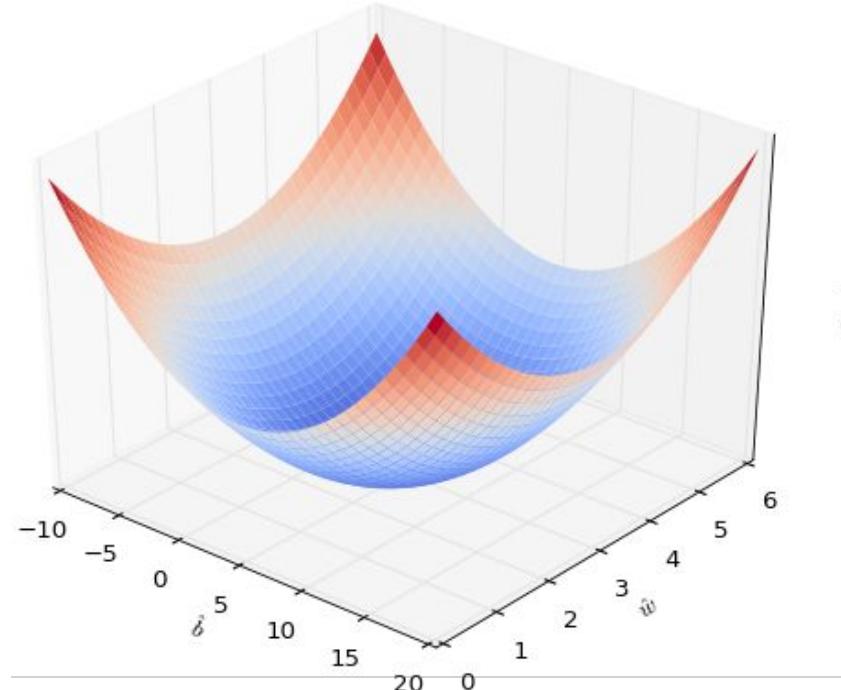
- Nem toda matriz tem uma matriz inversa

Desvantagens:

- Os atributos devem estar em escalas similares se não pode demorar muito para convergir (*StandardScaler*, ex.)
- Usa todo o conjunto de treinamento, o que deixa o algoritmo lento;

Solução

- Gradiente descendente estocástico ou Gradiente descendente mini-batch



Resumo

- Conheceu o algoritmo de Gradiente Descendente
- Entendeu as vantagens e desvantagens
- Conheceu outras soluções para o problema de otimização

Obrigado!

Você me encontra em:

- ◊ [@nickssonarrais](https://twitter.com/nickssonarrais)
- ◊ nickssonarrais@gmail.com

