# **Relatório Consolidado: Análise do Gradient Descent em Regressão Linear**

## **1. Objetivo Geral**

Avaliar o desempenho do algoritmo de Gradient Descent em regressão linear sob diferentes condições de pré-processamento e taxas de aprendizado (\u03b1). O estudo foi dividido em duas etapas:

* **Etapa 2:** Avaliação do impacto de diferentes taxas de aprendizado em dados sem normalização.
* **Etapa 3:** Investigação do efeito da normalização Min-Max na estabilidade e eficiência do algoritmo.

## **2. Etapa 2 – Experimentos com Taxas de Aprendizado**

### **2.1 Metodologia**

* **Dataset:** Relação entre população e lucro (ex1data1.txt).
* **Modelo:** Regressão Linear Univariada: .
* **Implementação:**
  + Função de custo: MSE (Mean Squared Error)
  + Gradient Descent vetorizado
  + Taxas de aprendizado testadas:
  + 1500 iterações por experimento

### **2.2 Resultados e Análises**

#### **Comportamento das Taxas de Aprendizado (Sem Normalização)**

|  | **Comportamento Observado** | **Custo Final** | **Observações** |
| --- | --- | --- | --- |
| 0.1 | Divergência imediata (NaN) | - | Overflow nas primeiras iterações |
| 0.03 | Oscilações violentas | Indefinido | Custo ultrapassa 1e10 |
| 0.01 | Convergência estável | ~4.48 | Curva suave |
| 0.003 | Convergência lenta | ~5.67 | Requer mais iterações |
| 0.001 | Quase nenhum progresso | ~29.85 | Passos muito pequenos |

### 

**Gráfico 1**: *Comparação de Convergência para Diferentes Taxas de Aprendizado (Sem Normalização)*

### **2.3 Observações Chave**

* **Taxas Altas ()**: Causaram instabilidade numérica devido à ausência de normalização.
* **Efeito da Normalização (antecipação):** Com dados normalizados, foi possível utilizar até 0.3 sem divergência.
* **Trade-off vs Iterações:** apresentou o melhor equilíbrio entre velocidade de convergência e estabilidade.

## **3. Etapa 3 – Impacto da Normalização Min-Max**

### **3.1 Metodologia**

* **Dataset:** Preço de imóveis com duas features:
  + Área construída (800–5000 sq-ft)
  + Número de quartos (1–5)
* **Casos Avaliados:**
  + Caso 1: Dados brutos (sem normalização)
  + Caso 2: Dados normalizados (Min-Max Scaling)
* **Parâmetros:**
  + Iterações: 400
  + (normalizado), (não normalizado)
* **Referência de Verificação:** Solução exata via Equação Normal

### **3.2 Resultados Principais**

#### **Estabilidade Numérica**

| Cenário | Divergência (NaN) | Máximo Permitido |
| --- | --- | --- |
| Sem Normalização | Sim |  |
| Com Normalização | Não | 0.3 |

#### **Desempenho de Convergência**

| Métrica | Sem Normalização | Com Normalização |
| --- | --- | --- |
| Iterações p/ | Não alcançado | 35 |
| Custo Final () | 2.1e+04 | 1.04e+04 |

### 

**Gráfico 2**: *Convergência do Gradient Descent com Diferentes Taxas de Aprendizado (Com Normalização)*

## **4. Conclusões Finais**

* A escolha adequada da taxa de aprendizado depende fortemente da escala dos dados.
* **Normalização Min-Max** é essencial para evitar instabilidade numérica e permitir taxas de aprendizado mais altas.
* Sem normalização, é necessário usar taxas extremamente pequenas, o que reduz drasticamente a eficiência.
* Para aplicações práticas, **normalizar sempre as features** antes de aplicar Gradient Descent é uma boa prática essencial.
* O valor de ideal, com dados normalizados, se encontra entre **0.03 e 0.3** para este tipo de problema.