**Nome: João Pedro de Alcântara Lima**

**Disciplina: EECP0053 - Tópicos Em Engenharia Da Computação II - Fundamentos de Redes Neurais.**

**Assunto/Atividade: Regressão Linear**

**1. Introdução**

A regressão linear é um dos métodos mais fundamentais em aprendizado de máquina. Sua eficiência, no entanto, depende de uma otimização eficaz dos parâmetros, comumente realizada por meio da técnica de descida do gradiente. Neste trabalho, exploramos visualmente a trajetória dos parâmetros durante a execução do algoritmo e sua convergência em relação à função de custo. Através dos experimentos conduzidos, também refletimos sobre a escolha adequada da taxa de aprendizado (α) e a inicialização dos parâmetros, conectando tais conceitos ao fine-tuning em redes neurais.

## 2. Resultados e Discussão

### 2.1 Trajetória da Descida do Gradiente

A Figura 1 mostra o contorno da função de custo J(θ), com as curvas de nível representando diferentes valores dessa função. Sobre o contorno, observa-se a trajetória do algoritmo de descida do gradiente (em vermelho), iniciando de um ponto arbitrário e caminhando em direção ao mínimo.

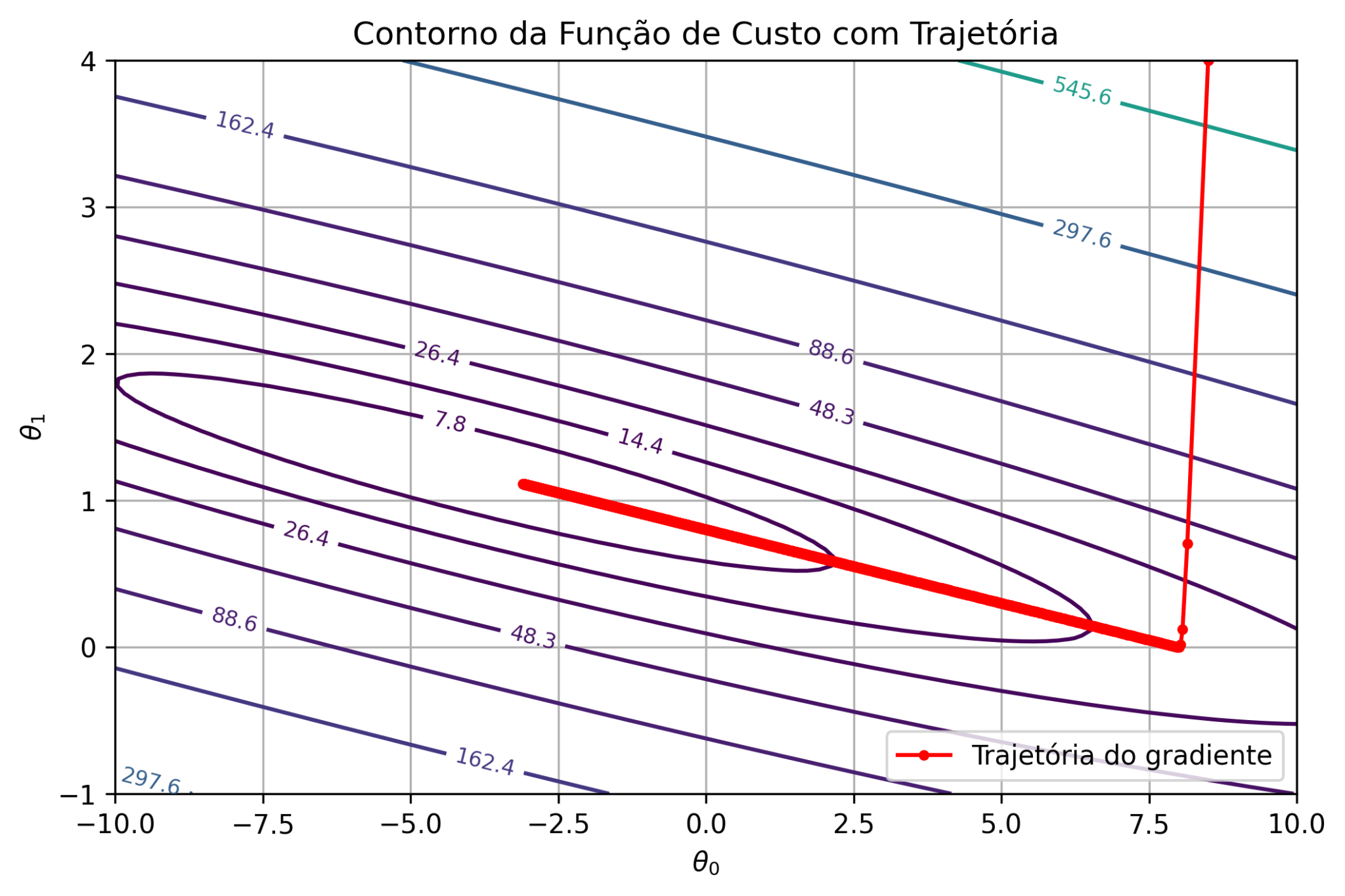


Figura 1 – Contorno da função de custo com trajetória da descida do gradiente.

### 2.2 Convergência da Função de Custo

A Figura 2 apresenta a evolução do valor da função de custo ao longo das iterações.  
Inicialmente, há uma queda brusca no custo, indicando que os primeiros passos trazem grandes ganhos. Com o passar das iterações, as melhorias tornam-se mais sutis, até que a função estabiliza próxima a um valor mínimo. Este comportamento é esperado em processos de otimização.

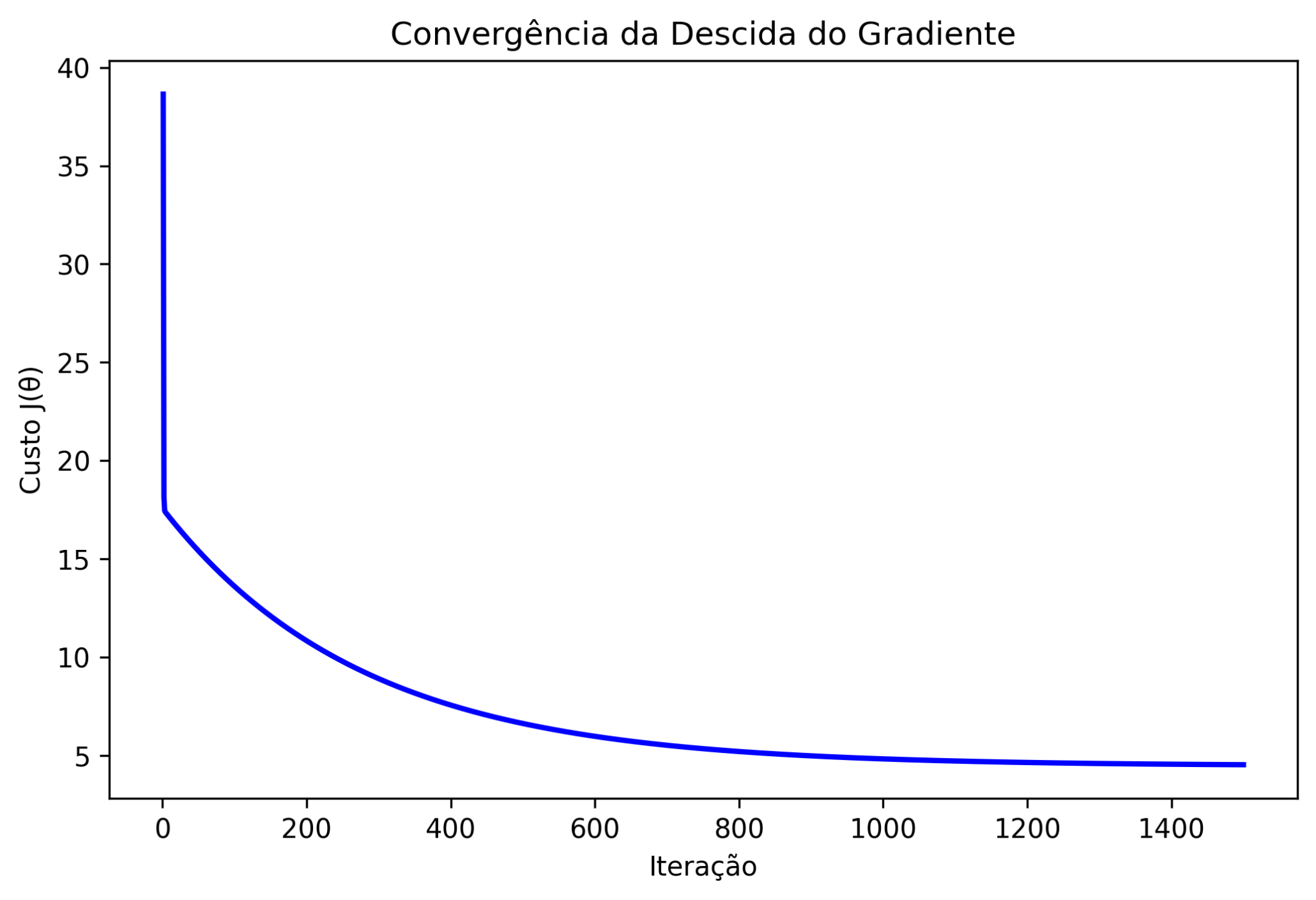


Figura 2 – Convergência da função de custo durante a descida do gradiente.

### 2.3 Visualização Tridimensional da Superfície de Custo

A Figura 3 mostra a superfície 3D da função de custo, permitindo observar como a trajetória da descida do gradiente ocorre no espaço tridimensional formado por θ₀, θ₁ e J(θ).  
A representação permite ver claramente o ponto inicial (em azul), a trajetória (em vermelho) e o ponto de convergência (em verde), ressaltando a eficiência do método na busca por mínimos locais.

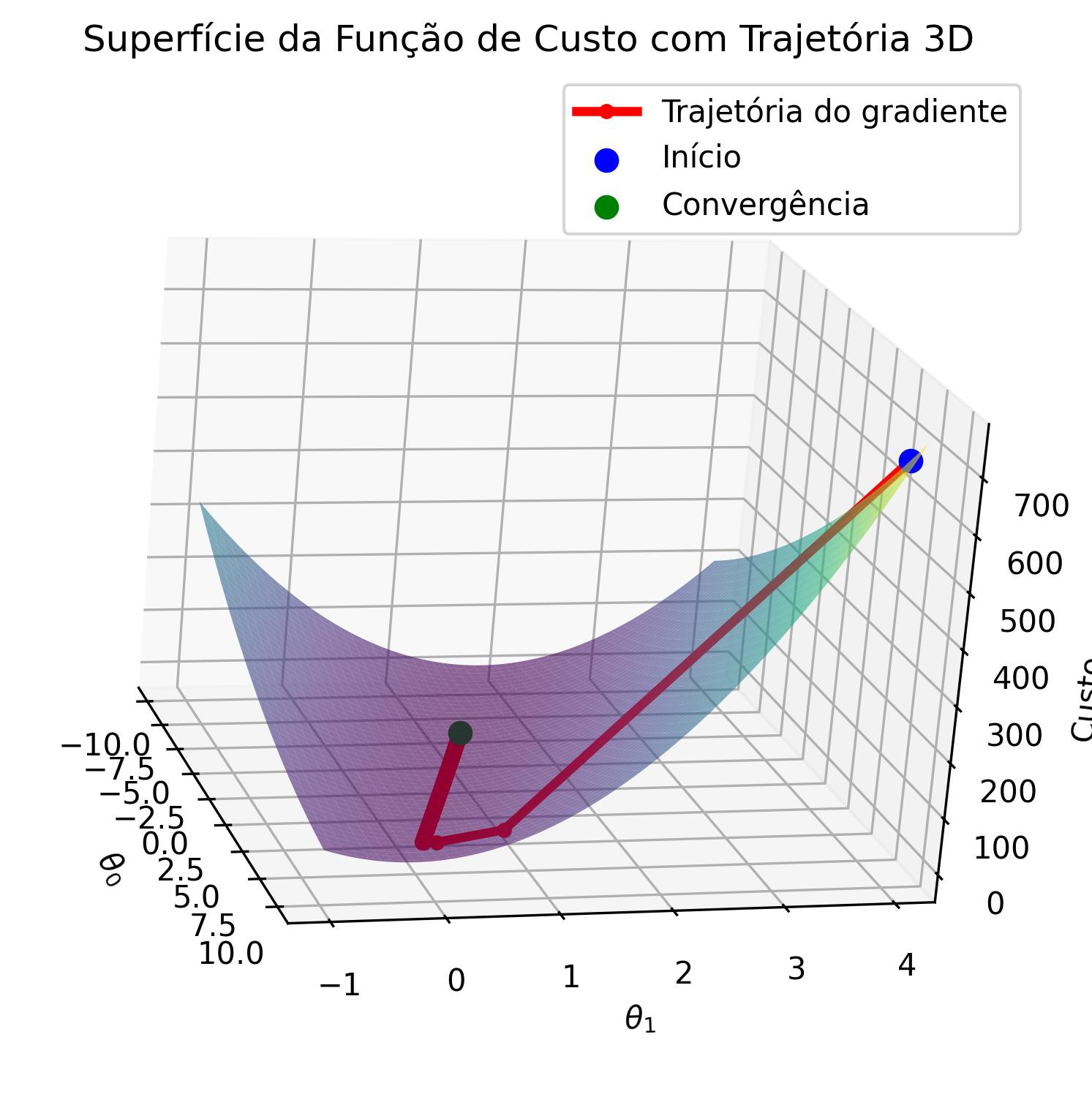


Figura 3 – Superfície da função de custo com trajetória da descida do gradiente.

### 2.4 Influência da Taxa de Aprendizado (α)

A escolha do valor de α é crítica. Quando α é muito grande, o algoritmo pode divergir, saltando sobre o mínimo e causando oscilações. Por outro lado, se α for muito pequeno, o algoritmo converge lentamente, aumentando o custo computacional. O valor ideal de α permite convergência rápida e estável, como demonstrado pelas figuras anteriores.

### 2.5 Inicialização dos Pesos e Fine-Tuning

A inicialização dos pesos (ou parâmetros θ) também influencia diretamente na eficiência da otimização. Uma inicialização inadequada pode levar a mínimos locais ou retardar a convergência. Em redes neurais profundas, é comum utilizar o chamado fine-tuning, que ajusta os pesos pré-treinados com uma pequena taxa de aprendizado, aproveitando uma boa inicialização para alcançar resultados mais precisos. Essa técnica demonstra a importância de começar a otimização a partir de um ponto já vantajoso, tal como ilustrado nas trajetórias observadas nas Figuras 1 e 3.

## 3. Conclusão

Através da visualização da função de custo e da trajetória da descida do gradiente, foi possível compreender de maneira clara a dinâmica do processo de otimização. A escolha da taxa de aprendizado e a inicialização dos pesos são fatores cruciais, não apenas para regressão linear, mas especialmente para métodos mais complexos, como redes neurais. A experimentação e análise visual permitiram uma compreensão mais profunda desses conceitos.