

Práctica 1: Regresión Lineal

Esta práctica aplicar el método de regresión lineal sobre unos datos de entrenamiento previamente dados.

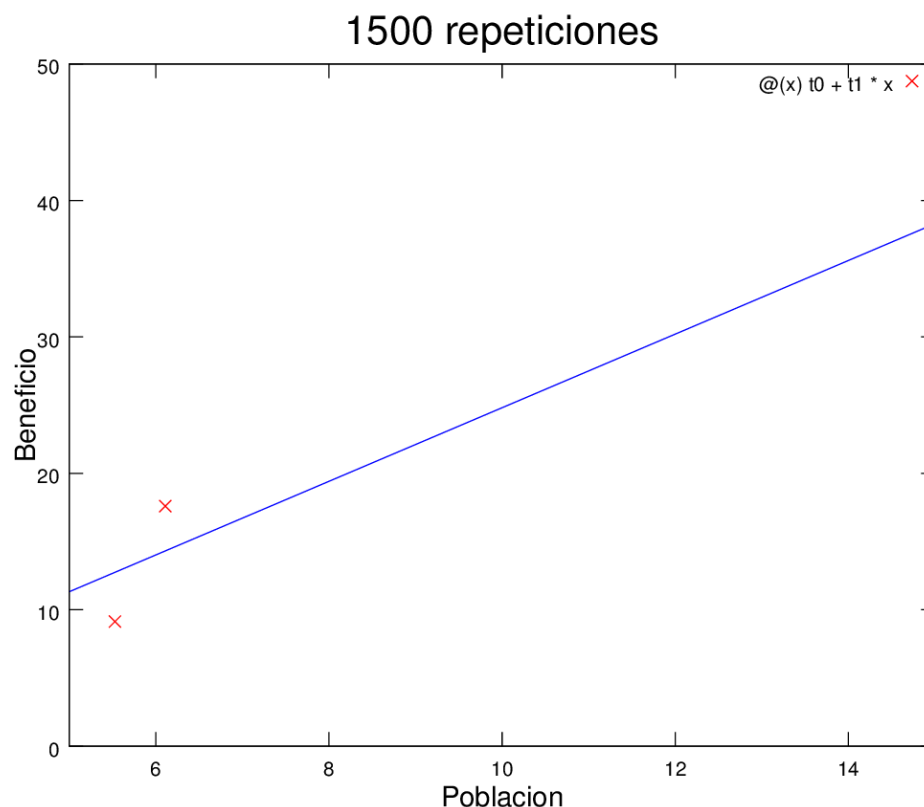
Hemos empleado 2 métodos de regresión:

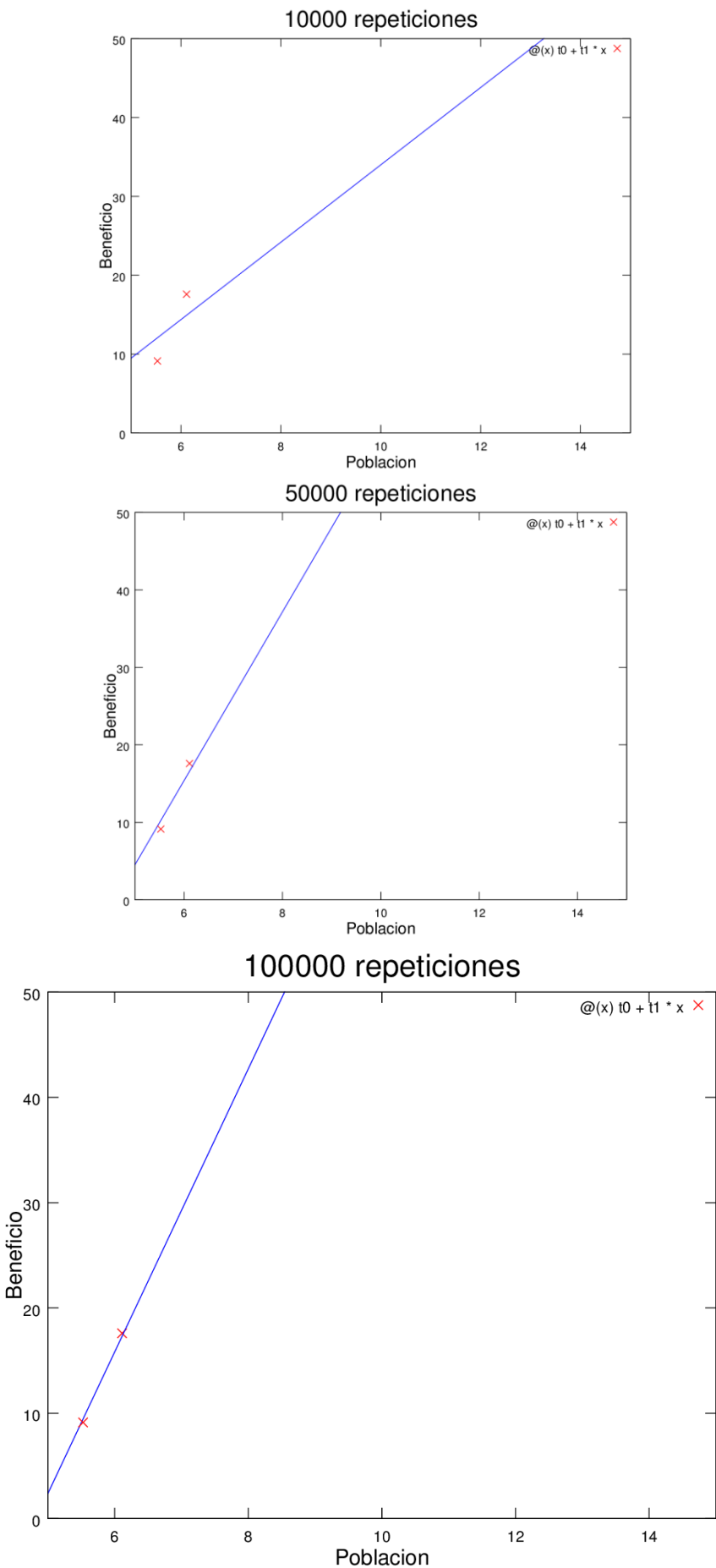
Regresión lineal con una variable

Hemos aplicado el método de descenso de gradiente para encontrar el θ que mejor se ajusta a los datos de entrenamiento, minimizando la función de coste J .

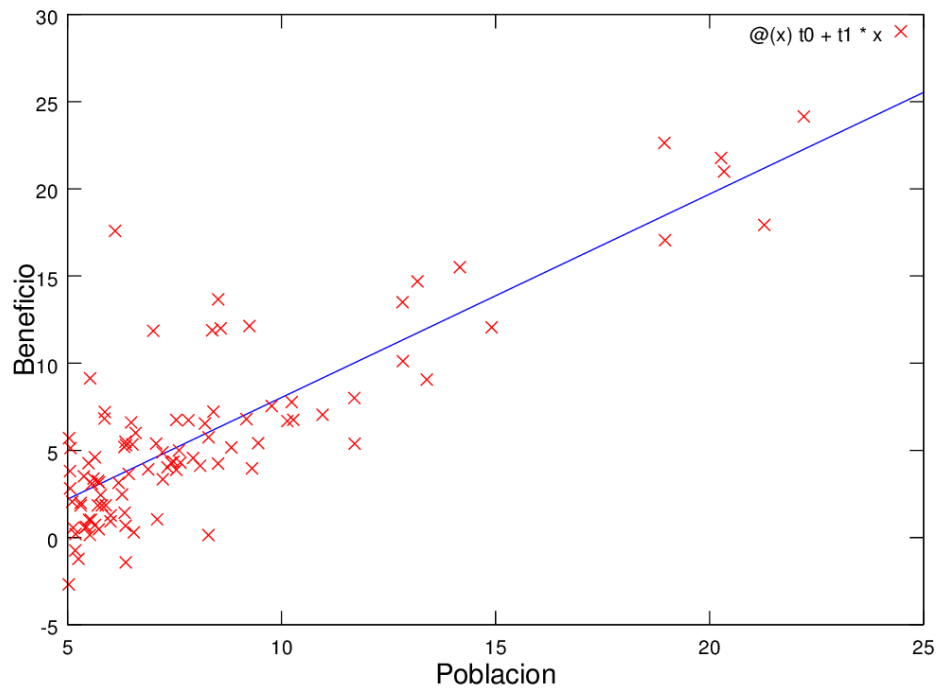
Para asegurarnos que hemos implementado bien el método del descenso de gradiente podemos hacer una prueba simple, que consiste en ver cómo se comporta la recta generada ante 2 únicos puntos de entrenamiento.

En las siguientes imágenes podemos comprobar la evolución de la misma.





Además podemos comprobar que los resultados tomando todos los datos de entrenamiento originales, son los correspondientes:



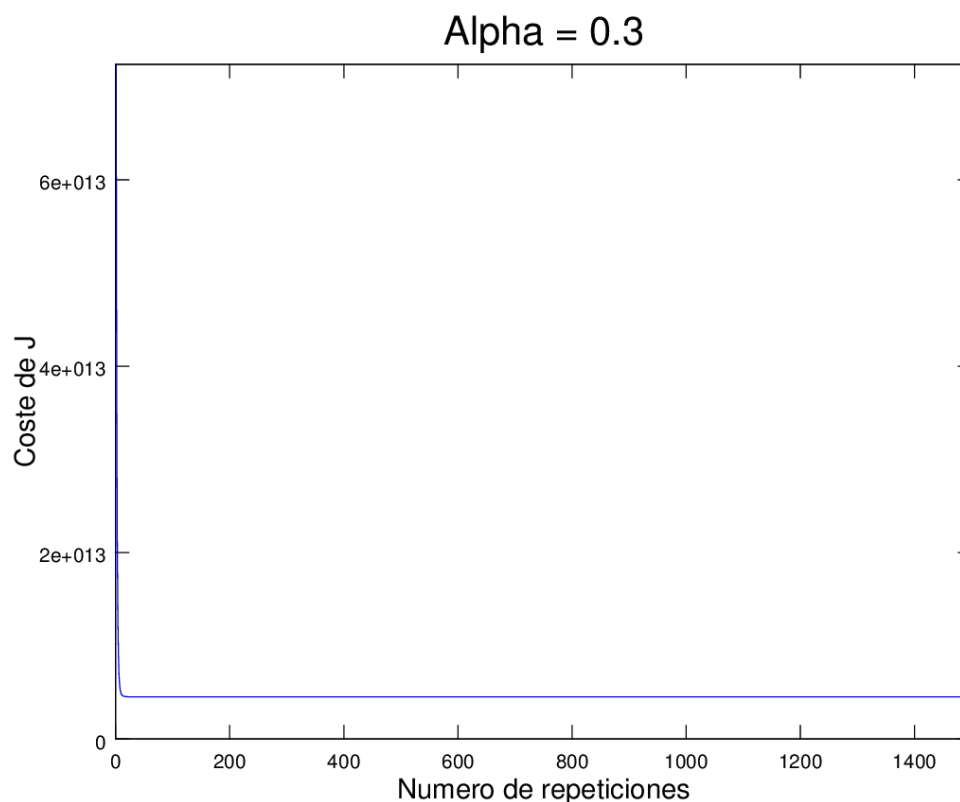
Regresión lineal con múltiples variables

Con estos segundos datos de entrenamiento hemos utilizado el método de regresión con varias variables porque se parten de varios atributos.

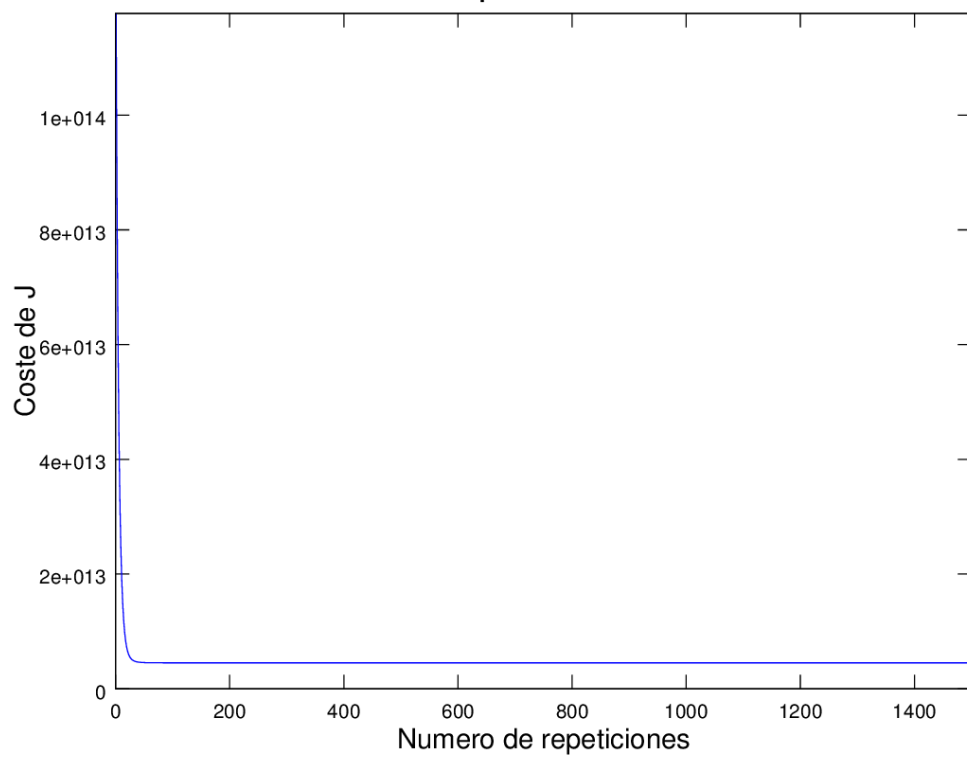
Como los datos parten de unidades muy diferentes, primero debemos normalizar los daos, aplicando una función de normalización.

Una vez normalizado, emplearemos el método de descenso de gradiente (similar al de una sola variable) y probaremos distintos valores de la tasa de aprendizaje alpha. Esto hace que se modifiquen los valores más lenta o rápidamente.

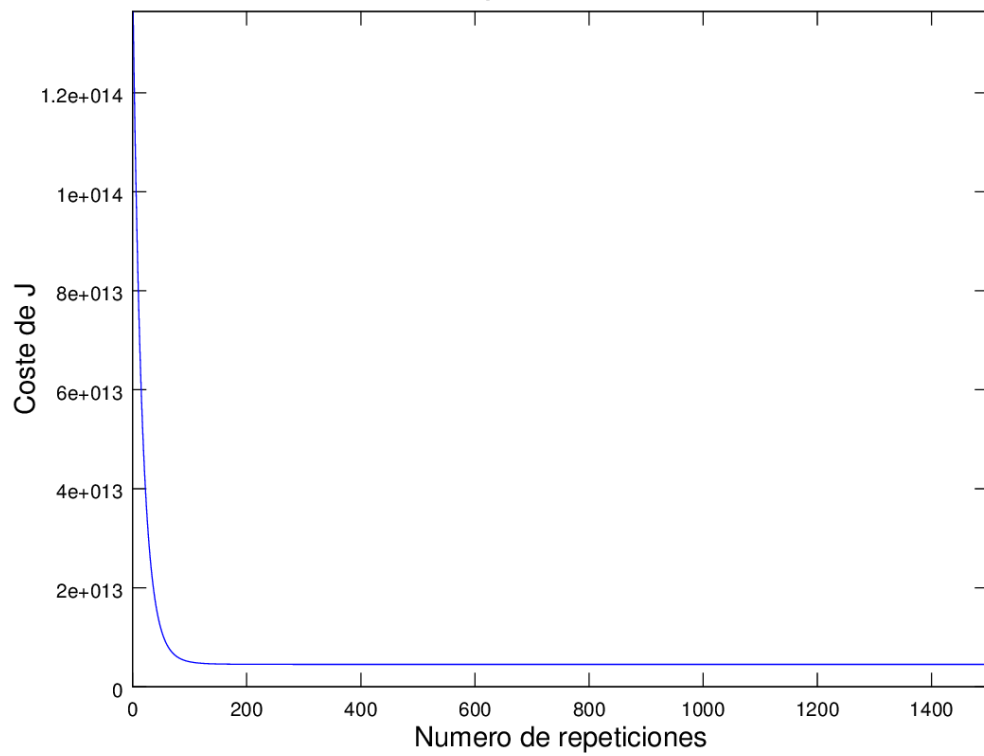
Las gráficas comparativas de cada valor de alpha son las siguientes:



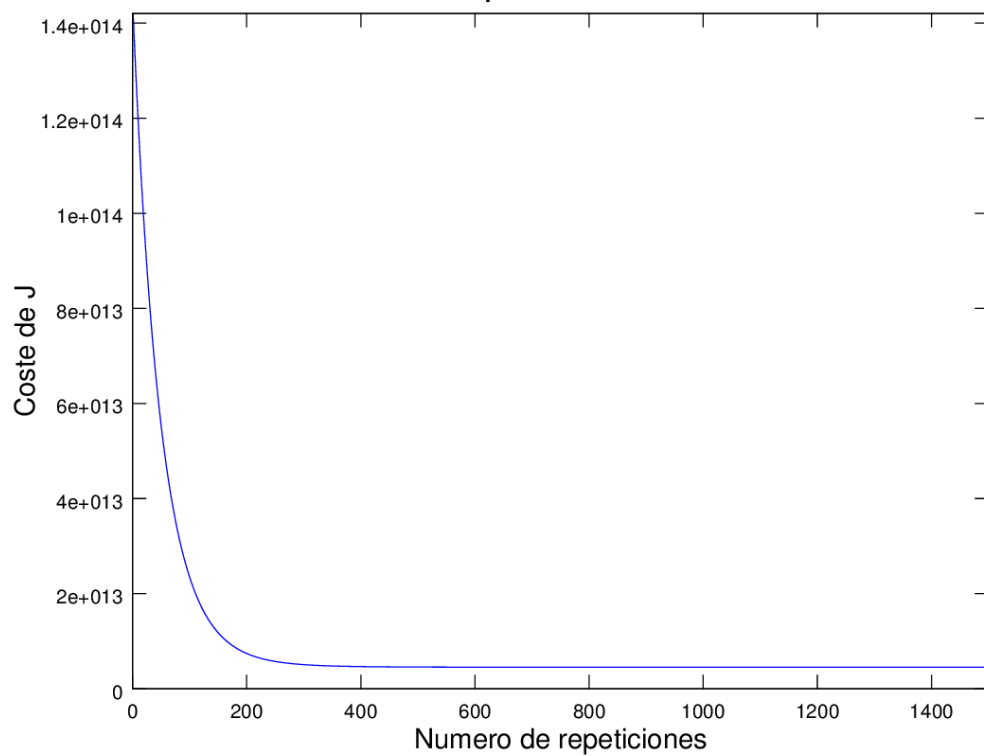
Alpha = 0.1



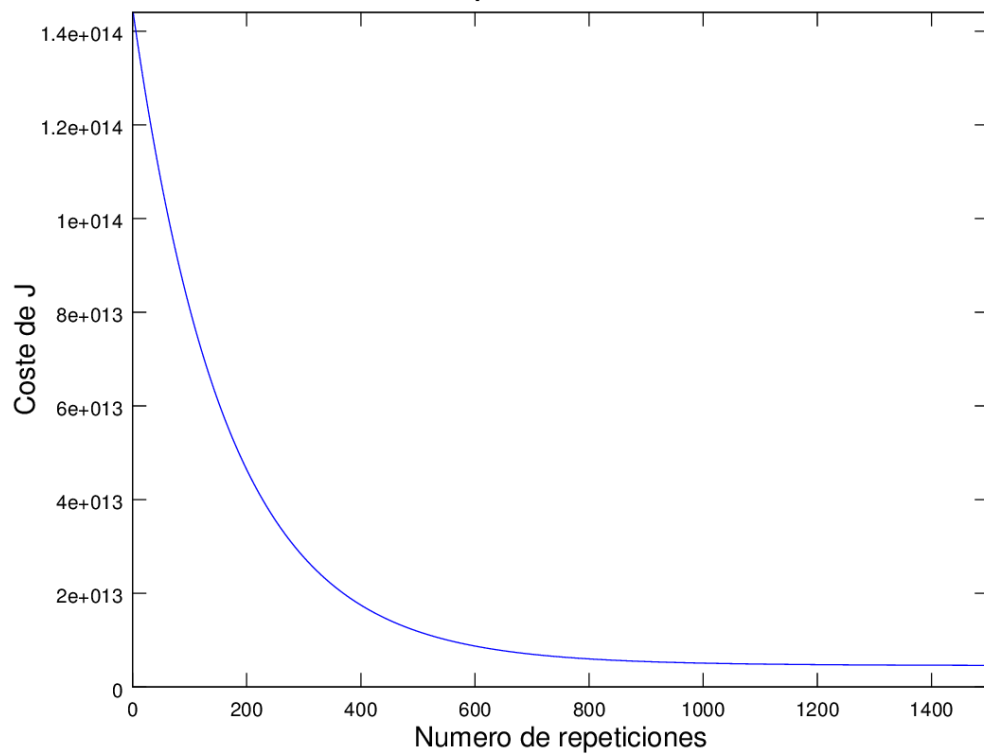
Alpha = 0.03

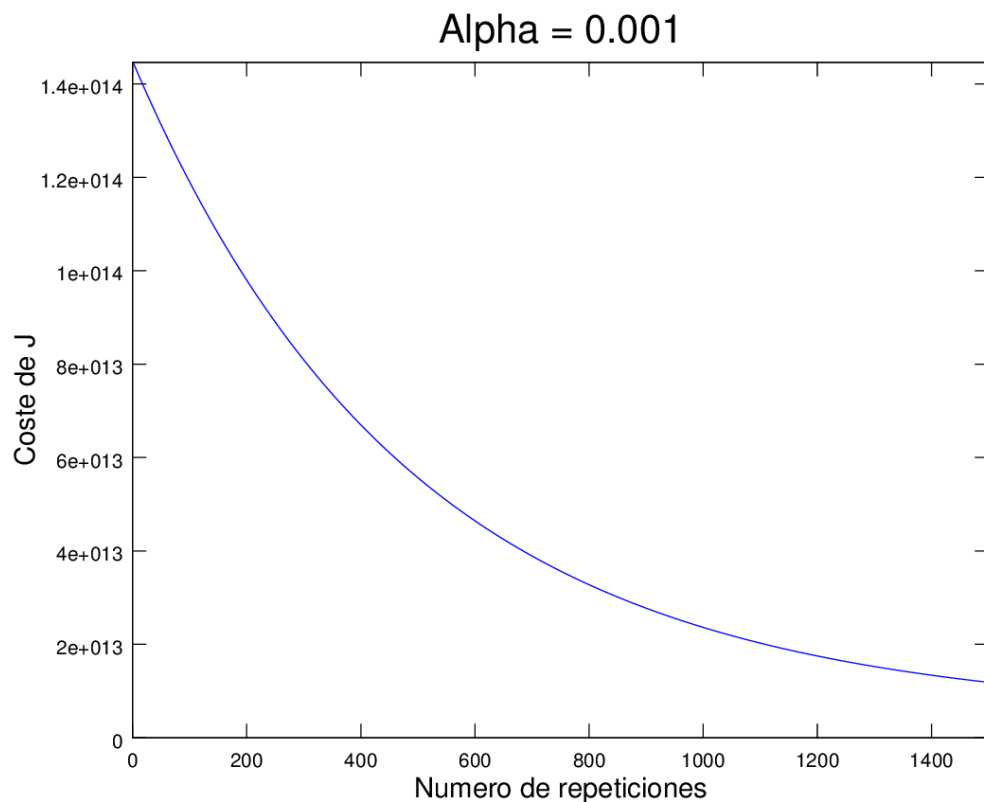


Alpha = 0.01



Alpha = 0.003





En estos ejemplos podemos apreciar que con un valor de alpha demasiado bajo, el número de iteraciones que se deben emplear para alcanzar el valor óptimo es exagerado, por lo que con las 1500 iteraciones no llegamos al mínimo.

Por último, realizaremos el mismo problema pero de una forma más sencilla, que es el método de la ecuación normal.

Además, comprobaremos que las estimaciones realizadas mediante ambos métodos (para unos mismos valores de entrada) son prácticamente iguales, con una pequeña desviación debido a que en la regresión quizá no se ha encontrado el valor más óptimo.