Introducción al Reconocimiento de Patrones

Tarea 3

Alexis Gavriel Gómez (2016085662)
Andrés Ramírez Quirós (2016142049)

Problema 1.1

1.1 $J(\theta) = \frac{1}{2} \Sigma\, w^i \,(\theta^T x^i - y^i)(\theta^T x^i - y^i)$

$= \frac{1}{2} \Sigma\, (\theta^T x^i - y^i)\, w^i\, (\theta^T x^i - y^i)$

Sea $(\theta^T x^i - y^i) = a_i$

$J(\theta) = \frac{1}{2} \sum_{i=1}^{m} a_i w^i a_i$

Se quiere que tenga la forma $x^T A x$ para que sea escalar si A es simetrica

Si $J(\theta) = \frac{1}{2} \sum_{i=1}^{m} a_i w^i a_i$

con $w^i a_i = b_i$

$J(\theta) = \sum_{i=1}^{m} a_i b_i = \frac{1}{2} \underline{A} \cdot \underline{B} = \frac{1}{2} A^T B$

Tomando en cuenta que $a_i = (\theta^T x^i - y^i)$

y $\Sigma\, \underline{\theta}^T x^{(i)} = X\theta$

$J(\theta) = \frac{1}{2} (X\theta - y)^T B$

$J(\theta) = \frac{1}{2} (X\theta - y)^T W (X\theta - y)$ con W una matriz diagonal

Problema 1.2

Ecuaciones normales X, W, Y

El mínimo sucede cuando $\nabla_\theta J(\theta) = 0$

$$\nabla_\theta J(\theta) = \nabla_\theta \tfrac{1}{2}(X\theta - y)^T W (X\theta - y) \overset{!}{=} 0$$

$$\tfrac{1}{2} \nabla_\theta (\theta^T X^T W X\theta - \theta^T X^T W y - y^T W X\theta + y^T W y) = 0$$

$$\tfrac{1}{2} \nabla_\theta (\theta^T X^T W X\theta - \theta^T X^T W y - y^T W X\theta) = 0$$

Como es un escalar real $tr\, A = A$

$$\tfrac{1}{2} \nabla_\theta (tr\, \theta^T X^T W X\theta - 2\, tr\, y^T W X\theta) = 0$$

Utilizando $\nabla_{A^T} tr\, ABA^T C = B^T A^T C^T + BA^T C$

$A = \theta^T$, $B = X^T W X$, $C = I$, $B^T = X^T W^T X^T$

Utilizando $\nabla_A tr\, AB = B^T$; $A = \theta$, $B = y^T W X$

$$\tfrac{1}{2}(X^T W X\theta + X^T W X\theta - 2 X^T W^T y) = 0$$

Como $W = W^T$

$$\tfrac{1}{2}(2 X^T W X\theta - 2 X^T W y) = 0$$

$$X^T W X\theta - X^T W y = 0$$

$$X^T W X\theta = X^T W y$$

$$\theta = (X^T W X)^{-1} X^T W y$$

Problema 1.3

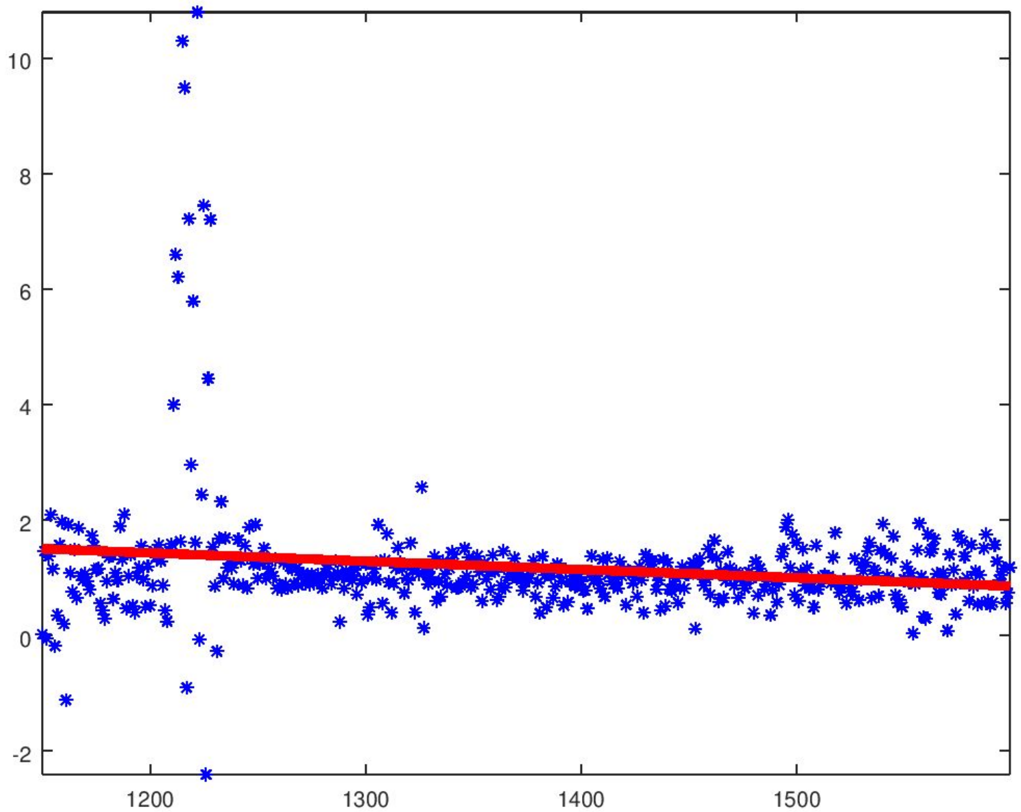
$$y = h_{\theta}(x) + \epsilon$$

$$\epsilon = J(\theta)$$

$$\ell(\theta) = m \ln \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{\theta^2} J(\theta)$$

$$\ell(\theta) = m \ln \frac{1}{\sqrt{2\pi}\sigma} - \frac{1}{\theta^2} \frac{1}{2} \sum (\theta^T x - y)(\theta^T x - y)$$

$$\text{con } \frac{1}{\theta_i^2} = w_i$$



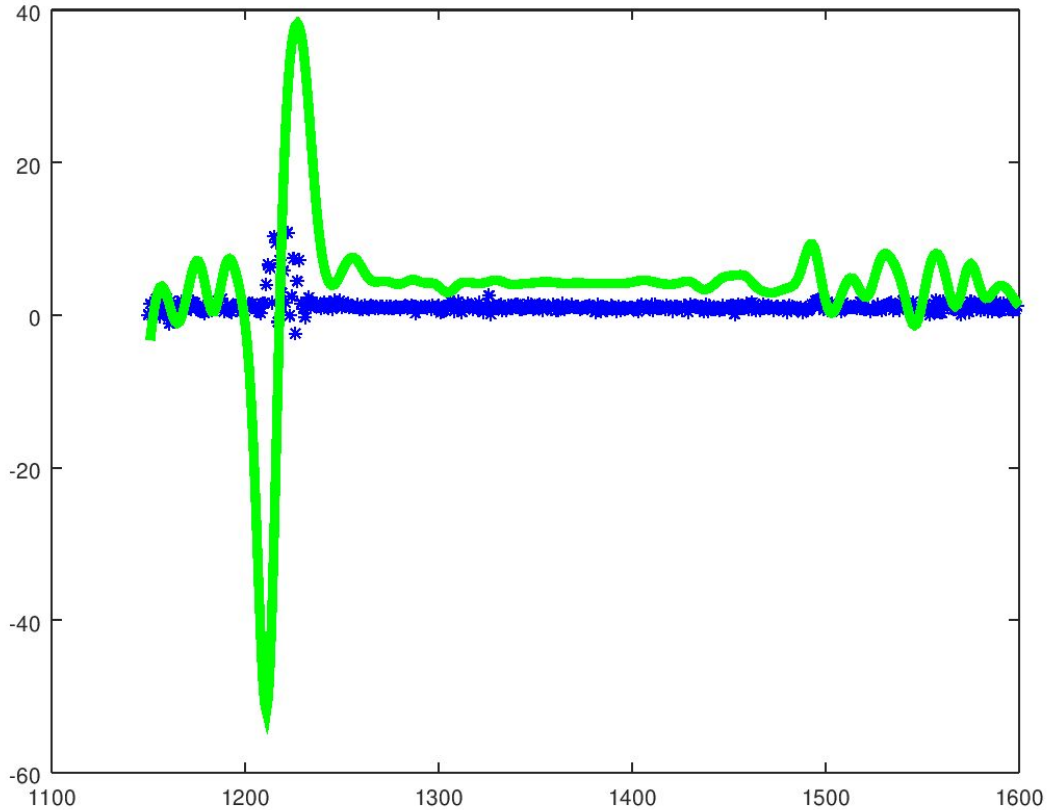
2.1

Para los datos de test_qso:

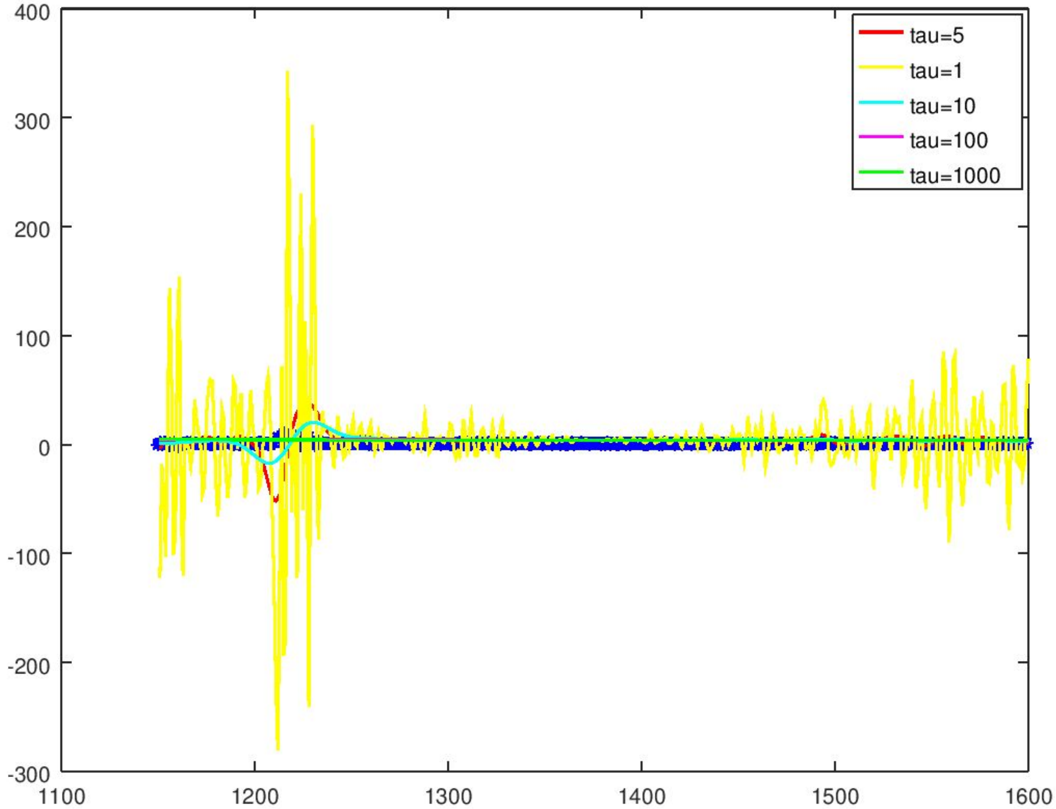
El theta óptimo resultante de la regresión lineal está dado por:

Theta = (-0.456575,-0.048912)

Problema 2.2



Problema 2.3



Tau representa el ancho de banda, por lo que entre mayor sea el número, más peso tendrán los valores lejanos al centro de la banda, esto causa que la linealización se comporte como un promedio.