

Homework on Linear and Logistic Regression

Machine Learning – 361-1-3761

מגיש: עומר לוכסמבורג

חלק 1 – רגרסיה לינארית

בחלק זה חישבנו את $f = C(\theta)$, ואת $g = \nabla_{\theta} C(\theta)$ והעברנו אותם לפונקציית מינימיזציה הנתונה.

מאחר ויש לנו 14 features ו-400 דוגמאות שונות, למעשה נקבל כי X הוא מטריצה בגודל 14 על 400.

לכן θ הוא וקטור בגודל 14 על 1. פונקציית המחיר שבחרתי היא פונקציית מחיר RMS , כאשר $n = 400$ והוא כמות הדגימות.

כלומר f הוא חישוב פונקציית המחיר הבאה:

$$C(\theta) = \frac{1}{2n} \sum_{i=1}^n (\theta^T x^{(i)} - y^{(i)})^2$$

ערך הגרדיאנט ישמר ב- g ויחושב כך:

$$\frac{\partial C(\theta)}{\partial \theta_j} = \frac{1}{2n} \sum_{i=1}^n 2(\theta^T x^{(i)} - y^{(i)}) \cdot x_j^{(i)}$$

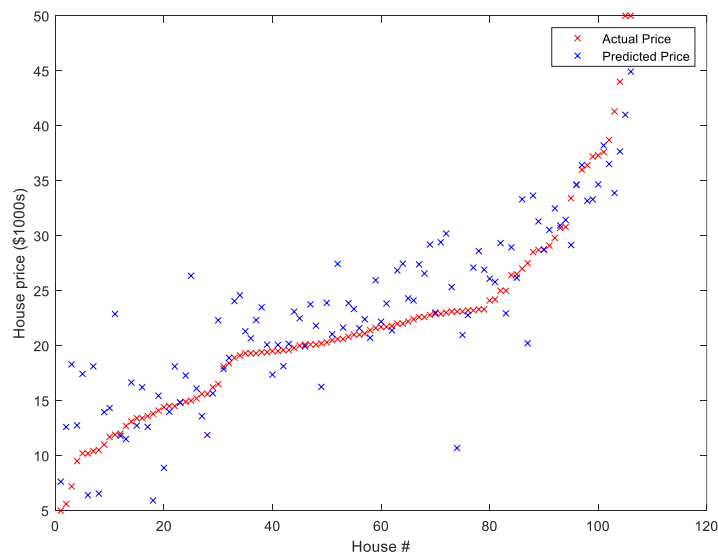
$$\theta_j^{t+1} = \theta_j^t - \frac{\alpha}{n} \sum_{i=1}^n (\theta^T x^{(i)} - y^{(i)}) \cdot x_j^{(i)}$$

↙ נמוץ פסאד

התוצאות:

```
Optimization took 0.224555 seconds.
RMS training error: 4.810984
RMS testing error: 4.253348
```

אכן קיבלנו ערך RMS בסביבות 4~5.



קוד מטלאב:

```
f = (1/(2*length(X))) * (norm(y - theta' * X,2))^2 ; % this is the RMS cost value
g = ((1/length(X)) * (theta'*X-y)*X')'; % this is the gradient calculation for theta
```

חלק 2 – Logistic Regression

בחלק זה ביצענו חישוב פונקציית מחיר וגרדיאנט עבור רגרסיה לוגיסטית כאשר השתמשנו בפונקצית סיגמוייד.

f הוא פונקציית המחיר המחושבת בדרך הבאה:

$$f = C(\theta) = -\frac{1}{n} \sum_{i=1}^n y^{(i)} \log(\sigma(\theta^T x^{(i)})) + (1 - y^{(i)}) \log(1 - \sigma(\theta^T x^{(i)}))$$

הגרדיאנט g מחושב כך:

$$g_j = \frac{\partial C(\theta)}{\partial \theta_j} = -\frac{1}{n} \sum_{i=1}^n (y^{(i)} - \sigma(\theta^T x^{(i)})) x_j^{(i)}$$

התוצאות:

```
Optimization took 2.516839 seconds.  
Training accuracy: 100.0%  
Test accuracy: 100.0%
```

קיבלנו כי 100% מה- *train* ומה- *test* זוהו נכון, ולכן ניתן להסיק כי החישובים נכונים.

קוד מטלאב:

```
f = -(1/length(X)) * (y*log(sigmoid(theta'*X))' + (1-y)*log(1-sigmoid(theta'*X))');  
g = -(1/length(X)) * ((y - sigmoid(theta'*X))*X)';
```