Problem Set # 1
Entrega: 10/Nov/2016 – 14:20 – em sala

Regras:

1. Os 10 problemas deverão ser solucionados individualmente.
2. Serão aceitos apenas trabalhos digitados.
3. Antes de buscar ajuda de colegas ou online, tente solucionar os problemas sozinho/a.
4. O aluno deverá citar as referências usadas (além do livro) e alunos com quem colaborou para solução do problema.
5. O trabalho será entregue em mãos no dia da aula.
6. Nem todas as questões serão corrigidas, mas todas deverão ser solucionadas para obter nota máxima.

Questões:

1. Problem 1.3
2. Problem 1.5
3. Problem 2.19
4. Problem 3.6
5. Problem 4.25
6. Problem 5.2
7. Elements of Statistical Learning (ESL 2nd ed) Exercise 2.7
8. ESL Exercise 4.2
9. ESL Exercise 3.27
10. (**Recursive Least Squares**) In this exercise we will derive the recursive least squares algorithm and then compare it with the (batch) least squares procedure.

    Consider the linear regression model

    $$y_i = x_i'\beta + \epsilon_i, \quad i = 1, \ldots, n$$

    where $x_1, ..., x_n$ are the regressors and $y_1, ..., y_n$ the dependent variables. We estimate the parameter vector β minimizing the sum pof squares, i.e.,

    $$\hat{\beta}_n = \arg\min_{\beta} \frac{1}{2n} \sum_{i=1}^{n} (y_i - \beta' x_i)^2 = \left(\sum_{i=1}^{n} x_i x_i' \right)^{-1} \sum_{i=1}^{n} x_i y_i = (\mathbf{X}_n' \mathbf{X}_n)^{-1} \mathbf{X}_n' \mathbf{y}_n,$$

    where $\mathbf{X}_n = (x_1, ..., x_n)'$ and $\mathbf{y} = (y_1, ..., y_n)'$.

    Define $P_k = (\mathbf{X}_k' \mathbf{X}_k)^{-1}$, we have $\hat{\beta}_k = P_k \mathbf{X}_k' \mathbf{y}_k$. In the next items we will show how to implement an iterative version of the least squares.

(a) Write the least squares estimator for $k+1$ observations using the decomposition $\mathbf{X}_{k+1} = (\mathbf{X}_k', x_{k+1})'$ and $\mathbf{y}_{n+1} = (\mathbf{y}_k', y_{k+1})'$. Your solution must only involve terms P_{k+1}, x_{k+1}, y_{k+1}, P_k, and $\hat{\beta}_k$.

(b) Show that $P_{k+1}^{-1} = P_n^{-1} + x_{k+1}x_{k+1}'$ and use this information to derive an interative formula for $\hat{\beta}_{k+1}$ in the form

$$\hat{\beta}_{k+1} = \hat{\beta}_k - A(b\hat{\beta}_k - c).$$

(c) Use the matrix inversion lemma:

$$\left(A - B'C^{-1}AB\right)^{-1} = A^{-1} + A^{-1}B'\left(C - BA^{-1}B'\right)^{-1}BA^{-1},$$

to derive an interative formula for $P_{k+1} = P_k + \nu_{k+1}R_{k+1}$, that does not require inverting a matrix.

(d) Write an algorithm based on (b)-(c) to estimate $\hat{\beta}_n$. Be mindful about the number of operations.

(e) Now let's model the TO BE DEFINED dataset using a linear regression model. Use the Recursive Least Squares algorithm and the Least Squares formula to estimate the parameters and compare their performance. Note that your implementation capability and the language used are crucial in the performance of both methods.