Machine Learning: Assignment 1

Sam Coxon

September 16, 2016

1. (a) Given data on past performances of certain soccer teams, we want to perform a supervised learning task to *classify* he outcome of an unknown future match

   (b)

	$Team1$	$Team2$	$Team3$
$Game1$	win	$loss$	win
$Game2$	$loss$	$draw$	win
$Game3$	win	$draw$	win

2. (a) To satisfy starting conditions, set $\theta_0 = 0$ and $\theta_1 = 1$ initially.
   Thus gradient descent algorithm for θ_0 is :

   $\theta_0 := \theta_0 - \alpha * \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)$

   $\theta_0 := 0 - (.1) * \frac{1}{3} * ((-3) + (-2) + (-4))$

   $= -(.1) * \frac{-9}{3}$

   $= .3$

   And gradient descent algorithm for θ_1 is :

   $\theta_1 := \theta_1 - \alpha * \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x_i) - y_i)x_i$

   $\theta_1 := 1 - (.1) * \frac{1}{3} * ((-3) * 3 + (-2) * 5 + (-4) * 6)$

   $= 1 - (.1) * \frac{-43}{3}$

   $= 2.4333$

Second iteration :

$\theta_0 := .3 - (.1) * \frac{1}{3} * ((.3 + 3 * 2.43 - 6) + (.3 + 5 * 2.43 - 7) + (.3 + 6 * 2.43 - 10))$

$= .3 - (.1) * \frac{1}{3} * 11.92$

$= -.09733$

$\theta_1 := 2.43 - (.1) * \frac{1}{3} * ((.3 + 3 * 2.43 - 6) * 3 + (.3 + 5 * 2.43 - 7) * 5 + (.3 + 6 * 2.43 - 10) * 6)$

$= 2.43 - (.1) * \frac{1}{3} * (4.77 + 27.25 + 29.28)$

$= .3899$

$MSE = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x) - y)^2$

$= \frac{1}{3}((-.097 + 3 * .398 - 6)^2 + (-.097 + 5 * .398 - 7)^2 + (-.097 + 6 * .398 - 10)^2)$

$= 36.5$

(b) Normalize data via $x_i = \frac{x_i - avg}{range}$

$x_1 = -.555$
$x_1 = .111$
$x_1 = .444$

$\theta_0 := 0 - (.1) * \frac{1}{3} * ((-.55 - 6) + (.11 - 7) + (.44 - 10))$

$\theta_0 := 0 - (.1) * \frac{1}{3} * (-23)$

$= .76$

$\theta_1 := 1 - (.1) * \frac{1}{3} * ((-6.55) * 3 + (-6.89) * 5 + (-9.56) * 6)$

$= 1 - (.1) * \frac{-111.46}{3}$

$= 4.72$

Second iteration :

$\theta_0 := .76-(.1)*\frac{1}{3}*((.76+-.55*4.72-6)+(.76+.11*4.72-7)+(.76+.44*4.72-10))$

$= .76 - (.1) * \frac{1}{3} * -20.72$

$= 1.45$

$\theta_1 := 4.72 - (.1) * \frac{1}{3} * ((.76 + -.55 * 4.72 - 6) * -.55 + (.76 + .11 * 4.72 - 7) * .11 + (.76 + .44 * 4.72 - 10) * .44)$

$= 4.72 - (.1) * \frac{1}{3} * (.53)$

$= 4.70$

4. When the partial derivative (w.r.t θ_1) of the cost function is equal to 0, this is the optimal θ_1 value. Thus:

   $\frac{1}{m} \sum_{i=1}^{m} (\theta_0 + \theta_1 * x_i - y_i) x_i = 0$

   isolated for θ_1 provides the optimal θ_1 value. So:

   $\frac{1}{m} * [((\theta_0 + \theta_1 * x_1 - y_1) * x_1 + (\theta_0 + \theta_1 * x_2 - y_2) * x_2 + ...] = 0$

   $= \frac{1}{m}[\sum_{i=1}^{m} \frac{(\theta_0 - y) * x}{x^2} + \sum_{i=1}^{m} \theta_1] = 0$

   $\theta_1 = \frac{1}{m}[\sum_{i=1}^{m} \frac{(\theta_0 - y)}{x}]$