机器学习第一讲习题以及课外阅读

> Tom M. Mitchell provided a widely quote: **"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E".**

习题

问题1. （任选两小题作答）

对给定样本$\{(X_i, Y_i)\}_{i=1,2,\ldots,100}$.**进行线性回归** $Y = \alpha X + \beta$.

(a) 线性回归是有监督学习还是无监督学习?简单解释理由.

(b) 描述吴恩达视频课程中梯度下降法求解的算法步骤.

(c) (**矩阵转置与乘法**)假定我们随机选取不同的样本子集并得到了3组不同的参数值$\{(\alpha_i, \beta_i)\}_{i=1,2,3.}$，令

$$\bar{X} = \begin{bmatrix} 1 & X_5 \\ 1 & X_{27} \\ 1 & X_{88} \end{bmatrix}, \bar{\theta} = \begin{bmatrix} \beta_1 & \alpha_1 \\ \beta_2 & \alpha_2 \\ \beta_3 & \alpha_3 \end{bmatrix}^T,$$

计算 $\bar{X}\bar{\theta}$.

课外阅读

1. 极大似然估计.

假定**问题1**中样本随机误差为独立同分步的正态随机变量，试用极大似然估计方法求解该线性回归问题.

2. 分位数线性回归损失函数的求导公式.

引理2.1(微积分基本定理): $F(b)-F(a)=\int_a^b F'(x)dx=\int_a^b dF(x)$
引理2.2: 假设 $g'(t)$是$[a,b]$上的连续函数; $f(t,x)$, 及其偏导数$f_x(t,x)$ 都是闭区域 $[a,b]\times[c,d]$上的连续函数, 且$\alpha(x),\beta(x)$是 $[c,d]$上的可微函数, 并满足 $a\le\alpha(x),\beta(x)\le b$, 则函数

$$F(x)=\int_{\alpha(x)}^{\beta(x)} f(t,x)dg(t)$$

在$[c,d]$上可微, 且有
$F'(x)=\int_{\alpha(x)}^{\beta(x)} f_x(t,x)dg(t)+f(\beta(x),x)g'(x)\beta'(x)-f(\alpha(x),x)g'(x)\alpha'(x).$

试求解 $G'(X)$, 已知 $G(X)=(\tau-1)\int_{-\infty}^{X}(x-X)dF(x)+\tau\int_X^{\infty}(x-X)dF(x)$, 其中 $F(\infty)=1,F(-\infty)=0.$