CS 229 Homework

Tyler Neylon

345.2016

These are solutions to the most recent problems posted for Stanford's CS 229 course, as of June 2016. I'm not sure if this course re-uses old problems, but please don't copy the answers if so.

1 Problem set 1

1.1 Logistic regression

1.1.1 Part (a)

The problem is to compute the Hessian matrix H for the function

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\log(g(y^{(i)}x^{(i)})),$$

where $g(z)$ is the logistic function, and to show that H is positive semi-definite; specifically, that $z^THz \geq 0$ for any vector z.

We'll use the fact that $g'(z) = g(z)(1-g(z))$. We'll also note that since all relevant operations are linear, it will suffice to ignore the summation over i in the definition of J. I'll use the notation ∂_j for $\frac{\partial}{\partial\theta_j}$, and introduce t for $y\theta^Tx$. Then

$$-\partial_j(mJ) = \frac{g(t)(1-g(t))}{g(t)}x_jy = x_jy(1-g(t)).$$

Next

$$-\partial_k\partial_j(mJ) = x_jy\big(-g(t)(1-g(t))\big)x_ky,$$

so that

$$\partial_{jk}(mJ) = x_j x_k y^2 \alpha,$$

where $\alpha = g(t)(1 - g(t)) > 0$.

Thus we can use repeated-index summation notation to arrive at

$$z^T H z = z_i h_{ij} z_j = (\alpha y^2)(z_i x_i x_j z_j) = (\alpha y^2)(x^T z)^2 \geq 0.$$

This completes this part of the problem.