Obligatory Assignment 1

Gabriel Sigurd Cabrera

STK-IN4300

October 14, 2019

Problem 2

We are given the *linearized* expression for the *object function*:

$$\mathbf{A} \equiv \sum_{i=1}^{N} g'(\mathbf{w}_{\text{old}}^{\mathrm{T}}\mathbf{x}_i)^2 \left(\frac{y_i - g(\mathbf{w}_{\text{old}}^{\mathrm{T}}\mathbf{x}_i)}{g'(\mathbf{w}_{\text{old}}^{\mathrm{T}}\mathbf{x}_i)} + \mathbf{w}_{\text{old}}^{\mathrm{T}}\mathbf{x}_i - \mathbf{w}^{\mathrm{T}}\mathbf{x}_i \right)^2 \tag{1}$$

We are interested in minimizing A; to accomplish this, we must take the derivative of A with respect to $\mathbf{w}$. We can then set this derivative to zero, and solve for $\mathbf{w}_{\text{min}}^{\mathrm{T}}$.

To accomplish this, we must redefine (1) such that its *summation notation* is replaced with a *vector/matrix* expression; we begin by redefining some terms. Let $p_i \equiv \frac{y_i - g(\mathbf{w}_{\text{old}}^{\mathrm{T}}\mathbf{x}_i)}{g'(\mathbf{w}_{\text{old}}^{\mathrm{T}}\mathbf{x}_i)} + \mathbf{w}_{\text{old}}^{\mathrm{T}}\mathbf{x}_i$, and let $q_i \equiv g'(\mathbf{w}_{\text{old}}^{\mathrm{T}}\mathbf{x}_i)^2$. This gives us:

$$= \sum_{i=1}^{N} q_i (p_i - \mathbf{w}^{\mathrm{T}}\mathbf{x}_i)^2 = (\mathbf{p} - \mathbf{w}^{\mathrm{T}}\mathbf{x}^{\mathrm{T}})^{\mathrm{T}}\mathbf{q}(\mathbf{p} - \mathbf{w}^{\mathrm{T}}\mathbf{x}^{\mathrm{T}})$$

We then expand the above, giving us several easily-differentiable terms:

$$= (\mathbf{p}^{\mathrm{T}}\mathbf{q} - \mathbf{wxq})(\mathbf{p} - \mathbf{w}^{\mathrm{T}}\mathbf{x}^{\mathrm{T}}) = \mathbf{p}^{\mathrm{T}}\mathbf{qp} - \mathbf{wxqp} - \mathbf{p}^{\mathrm{T}}\mathbf{qw}^{\mathrm{T}}\mathbf{x}^{\mathrm{T}} + \mathbf{wxqw}^{\mathrm{T}}\mathbf{x}^{\mathrm{T}}$$

We can then differentiate the above with respect to $\mathbf{w}$:

$$\frac{\partial \mathbf{A}}{\partial \mathbf{w}} = -\mathbf{xqp} - \mathbf{p}^{\mathrm{T}}\mathbf{qx}^{\mathrm{T}} + 2\mathbf{xqw}^{\mathrm{T}}\mathbf{x}^{\mathrm{T}} = -2\mathbf{xqp} + 2\mathbf{xqw}^{\mathrm{T}}\mathbf{x}^{\mathrm{T}}$$