Machine Learning – Brett Bernstein

Week 1 Lab: Concept Check Exercises

Starred problems are optional.

Multivariable Calculus Exercises

1. If $f'(x;u) < 0$ show that $f(x+hu) < f(x)$ for sufficiently small $h > 0$.
2. Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable, and assume that $\nabla f(x) \neq 0$. Prove
$$\arg\max_{\|u\|_2=1} f'(x;u) = \frac{\nabla f(x)}{\|\nabla f(x)\|_2} \quad \text{and} \quad \arg\min_{\|u\|_2=1} f'(x;u) = -\frac{\nabla f(x)}{\|\nabla f(x)\|_2}.$$
3. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x,y) = x^2 + 4xy + 3y^2$. Compute the gradient $\nabla f(x,y)$.
4. Compute the gradient of $f : \mathbb{R}^n \to \mathbb{R}$ where $f(x) = x^T Ax$ and $A \in \mathbb{R}^{n\times n}$ is any matrix.
5. Compute the gradient of the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ given by
$$f(x) = b + c^T x + x^T Ax,$$
where $b \in \mathbb{R}$, $c \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$.
6. Fix $s \in \mathbb{R}^n$ and consider $f(x) = (x-s)^T A(x-s)$ where $A \in \mathbb{R}^{n\times n}$. Compute the gradient of f.
7. Consider the ridge regression objective function
$$f(w) = \|Aw - y\|_2^2 + \lambda\|w\|_2^2,$$
where $w \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $y \in \mathbb{R}^m$, and $\lambda \in \mathbb{R}_{\geq 0}$.
   (a) Compute the gradient of f.
   (b) Express f in the form $f(w) = \|Bw - z\|_2^2$ for some choice of B, z.
   (c) Using either of the parts above, compute
$$\arg\min_{w\in\mathbb{R}^n} f(w).$$
8. Compute the gradient of
$$f(\theta) = \lambda\|\theta\|_2^2 + \sum_{i=1}^{n} \log(1 + \exp(-y_i\theta^T x_i)),$$
where $y_i \in \mathbb{R}$ and $\theta \in \mathbb{R}^m$ and $x_i \in \mathbb{R}^m$ for $i = 1,\dots,n$.

Linear Algebra Exercises

1. When performing linear regression we obtain the *normal equations* $A^T Ax = A^T y$ where $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$, and $y \in \mathbb{R}^m$.

   (a) If $\textbf{rank}(A) = n$ then solve the normal equations for x.

   (b) $(\star)$ What if $\textbf{rank}(A) \neq n$?

2. Prove that $A^T A + \lambda \mathbf{I}_{n\times n}$ is invertible if $\lambda > 0$ and $A \in \mathbb{R}^{n\times n}$.

3. $(\star)$ Describe the following set geometrically:

$$\left\{ v \in \mathbb{R}^2 \;\middle|\; v^T \begin{pmatrix} 2 & 2 \\ 0 & 2 \end{pmatrix} v = 4 \right\}.$$