

# CS 229, Fall 2018

## Problem Set #0: Linear Algebra and Multivariable Calculus

**Notes:** (1) These questions require thought, but do not require long answers. Please be as concise as possible. (2) If you have a question about this homework, we encourage you to post your question on our Piazza forum, at <https://piazza.com/stanford/fall2018/cs229> (3) If you missed the first lecture or are unfamiliar with the collaboration or honor code policy, please read the policy on Handout #1 (available from the course website) before starting work. (4) This specific homework is *not graded*, but we encourage you to solve each of the problems to brush up on your linear algebra. Some of them may even be useful for subsequent problem sets. It also serves as your introduction to using Gradescope for submissions.

### 1. [0 points] Gradients and Hessians

Recall that a matrix  $A \in \mathbb{R}^{n \times n}$  is *symmetric* if  $A^T = A$ , that is,  $A_{ij} = A_{ji}$  for all  $i, j$ . Also recall the gradient  $\nabla f(x)$  of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , which is the  $n$ -vector of partial derivatives

$$\nabla f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{bmatrix} \quad \text{where } x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

The hessian  $\nabla^2 f(x)$  of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is the  $n \times n$  symmetric matrix of twice partial derivatives,

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} f(x) & \frac{\partial^2}{\partial x_1 \partial x_2} f(x) & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(x) & \frac{\partial^2}{\partial x_2^2} f(x) & \cdots & \frac{\partial^2}{\partial x_2 \partial x_n} f(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x) & \frac{\partial^2}{\partial x_n \partial x_2} f(x) & \cdots & \frac{\partial^2}{\partial x_n^2} f(x) \end{bmatrix}.$$

- (a) Let  $f(x) = \frac{1}{2}x^T A x + b^T x$ , where  $A$  is a symmetric matrix and  $b \in \mathbb{R}^n$  is a vector. What is  $\nabla f(x)$ ?
- (b) Let  $f(x) = g(h(x))$ , where  $g : \mathbb{R} \rightarrow \mathbb{R}$  is differentiable and  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  is differentiable. What is  $\nabla f(x)$ ?
- (c) Let  $f(x) = \frac{1}{2}x^T A x + b^T x$ , where  $A$  is symmetric and  $b \in \mathbb{R}^n$  is a vector. What is  $\nabla^2 f(x)$ ?
- (d) Let  $f(x) = g(a^T x)$ , where  $g : \mathbb{R} \rightarrow \mathbb{R}$  is continuously differentiable and  $a \in \mathbb{R}^n$  is a vector. What are  $\nabla f(x)$  and  $\nabla^2 f(x)$ ? (*Hint:* your expression for  $\nabla^2 f(x)$  may have as few as 11 symbols, including ' and parentheses.)