Derivation of Hessian of log-sum-exp function

Jiayu Ye

August 1, 2018

1 Definition

The function $f(x) = \log(e^{x_1} + e^{x_2} + \cdots + e^{x_n})$ is called log-sum-exp function. The hessian of $f(x)$ is

$$\nabla^2 f(x) = \frac{1}{(1^T z)^2}((1^T z)\mathrm{diag}(z) - zz^T)$$

where $z = (e^{x_1}, \cdots, e^{x_n})$. [1]

2 Proof (vector calculus)

First denote $x = (x_1, \cdots, x_n)$, a.k.a a row vector as z.

$$\nabla^2 f(x) = \frac{\partial}{\partial x^T} \frac{\partial f(x)}{\partial x} \tag{1}$$

$$\frac{\partial f(x)}{\partial x} = \frac{\partial \log(1^T z)}{\partial x} = \frac{1}{1^T z} \frac{\partial (1^T z)}{\partial x} = \frac{1}{1^T z} z$$

$$\frac{\partial}{\partial x^T}\left(\frac{z}{1^T z}\right) = \frac{\frac{\partial z}{\partial x^T} 1^T z - z \frac{\partial (1^T z)}{\partial x^T}}{(1^T z)^2} = \frac{(1^T z)\mathrm{diag}(z) - zz^T}{(1^T z)^2}$$

3 Vector Calculus

Explanation over (1):
We assume $\nabla f(x)$ produces a row vector. (Notice that producing a column vector would yield the same result).

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

[1] Convex Optimization – Boyd and Vandenberghe Page 74

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

Consider the element $\nabla(\frac{\partial f}{\partial x_1}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} \\ \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} \end{bmatrix}$, hence $\mathbf{H} = \frac{\partial}{\partial x^T} \frac{\partial f(x)}{\partial x}$.