1. Given $f(x) = 0.5x_1^2 + \sin x_2 + e^{x_3}$ where $x \in \mathbb{R}^3$, which of the following is $\partial f/\partial x$? (hint: $f : \mathbb{R}^3 \to \mathbb{R}$, think about the dimensions)

   (a) $\frac{\partial f}{\partial x} = \begin{bmatrix} x_1 \\ \cos x_2 \\ e^{x_3} \end{bmatrix} \in \mathbb{R}^{3\times 1}$

   (b) $\frac{\partial f}{\partial x} = \begin{bmatrix} x_1 & \cos x_2 & e^{x_3} \end{bmatrix} \in \mathbb{R}^{1\times 3}$

   Solution: (b), the Jacobian $\partial f/\partial x$ will have the same number of rows as outputs to the function, and same number of columns as inputs to the function.

2. Given $f(x) = 0.5x_1^2 + \sin x_2 + e^{x_3}$ where $x \in \mathbb{R}^3$, which of the following is $\nabla_x f(x)$? (hint: $f : \mathbb{R}^3 \to \mathbb{R}$, think about the dimensions)

   (a) $\nabla_x f(x) = \begin{bmatrix} x_1 \\ \cos x_2 \\ e^{x_3} \end{bmatrix} \in \mathbb{R}^{3\times 1}$

   (b) $\nabla_x f(x) = \begin{bmatrix} x_1 & \cos x_2 & e^{x_3} \end{bmatrix} \in \mathbb{R}^{1\times 3}$

   Solution: (a), the transpose of a 1D Jacobian is known as the gradient ($\nabla_x f(x)$) and it is the same size as the function's input.

3. Given $f(x) = 0.5x_1^2 + \sin x_2 + e^{x_3}$ where $x \in \mathbb{R}^3$, is $\frac{\partial}{\partial x}(\nabla_x f(x)) = \nabla_x^2 f(x)$?

   (a) yes

   (b) no

   Solution: (a). The Jacobian of the gradient of a function is the hessian of that function.

4. Which one of the following pairs of vectors $x \in \mathbb{R}^3$ and $y \in \mathbb{R}^3$ **do not** exhibit complementarity?

   (a) $x = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad y = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

   (b) $x = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad y = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

   (c) $x = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

   Solution: (c), since the element-wise product (Hadamard product) of (a) and (b) is $[0, 0, 0]^T$, but for (c) it's $[1, 1, 0]^T$.

5. If we have the following optimization problem:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & c(x) = 0 \end{array} \tag{1}$$

   with a dual variable λ associated with the constraint $c(x) = 0$ and a Lagrangian $\mathcal{L}(x, \lambda) = f(x) + \lambda^T c(x)$, we have the following KKT conditions:

$$\nabla_x \mathcal{L}(x, \lambda) = \nabla_x f(x) + \left[\frac{\partial c}{\partial x}\right]^T \lambda = 0, \tag{2}$$

$$c(x) = 0. \tag{3}$$

   Which one of the following linear systems computes the "full" Newton step, (the other being the Gauss-Newton step)?

(a) $\begin{bmatrix} \nabla_x^2 f(x) & \left[\frac{\partial c}{\partial x}\right]^T \\ \frac{\partial c}{\partial x} & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -\nabla_x \mathcal{L}(x, \lambda) \\ -c(x) \end{bmatrix}$

(b) $\begin{bmatrix} \nabla_x^2 \mathcal{L}(x, \lambda) & \left[\frac{\partial c}{\partial x}\right]^T \\ \frac{\partial c}{\partial x} & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} -\nabla_x \mathcal{L}(x, \lambda) \\ -c(x) \end{bmatrix}$

Solution: (b). The only difference between the full Newton step and the Gauss Newton step is the upper left block of the KKT Jacobian. In Full Newton, this is the hessian of the Lagrangian, and in the Gauss-Newton case this is approximated as the hessian of the cost function.