1) Consider different **linear** regression problems, for which we have defined design matrices X_1, X_2, X_3. For each design matrix, write down the corresponding hypothesis $h_\theta(x)$.

$$X_1 = \begin{array}{c} \begin{array}{cc} x_0 & x \end{array} \\ \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \end{array}, \quad X_2 = \begin{array}{c} \begin{array}{ccc} x_0 & x & x^2 \end{array} \\ \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix} \end{array}, \quad X_3 = \begin{array}{c} \begin{array}{cc} x & x^2 \end{array} \\ \begin{bmatrix} 2 & 4 \\ 3 & 9 \\ 4 & 16 \end{bmatrix} \end{array}$$

2) Consider different **logistic** regression problems, for which we have defined design matrices X_1, X_2, X_3. For each design matrix, write down the corresponding hypothesis $h_\theta(x)$.

$$X_1 = \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{bmatrix} 1 & 3 & 5 \\ 1 & 5 & 11 \\ 1 & 7 & 16 \end{bmatrix} \end{array}, \quad X_2 = \begin{array}{c} \begin{array}{ccc} x_0 & x & x^2 \end{array} \\ \begin{bmatrix} 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 6 & 36 \end{bmatrix} \end{array}, \quad X_3 = \begin{array}{c} \begin{array}{cc} x & \cos(x) \end{array} \\ \begin{bmatrix} 0 & 1 \\ \frac{\pi}{2} & 0 \\ \frac{\pi}{4} & \frac{\sqrt{2}}{2} \end{bmatrix} \end{array}$$

3) The logistic regression hypothesis function is given by $h_\theta(x) = \sigma(x^T\theta) = \sigma(\theta_0 x_0 + \theta_1 x_1 + \ldots + \theta_n x_n)$ with the sigmoid function $\sigma(z) = \frac{1}{1+e^{-z}}$, parameters $\theta = (\theta_0, \theta_1, \ldots, \theta_n)^T$, and input vector $x = (x_0, x_1, \ldots, x_n)^T$. By convention the constant feature x_0 is fixed to $x_0 = 1$. With $x^{(i)} = (x_0^{(i)}, x_1^{(i)}, \ldots, x_n^{(i)})^T,\ x_0^{(i)} = 1$ and $\log(\,\cdot\,)$ referring to the natural logarithm, the logistic regression cost function can be written as

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left(y^{(i)}\log(h_\theta(x^{(i)})) + (1-y^{(i)})\log(1-h_\theta(x^{(i)}))\right).$$

Derive the gradient of the cost function, i.e. show that the partial derivative of the cost function with respect to θ_j equals

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)\cdot x_j^{(i)}.$$

Hint: Note that the derivative of the sigmoid function σ verifies $\dfrac{\partial \sigma(z)}{\partial z} = \sigma(z)\cdot(1-\sigma(z))$.