A Note on Gradients

In case you haven't seen it before, if $f : \mathbb{R}^n \to \mathbb{R}$ is a function that takes in n real numbers and outputs a single real number, its *gradient*, written ∇f, is a function $f : \mathbb{R}^n \to \mathbb{R}^n$. Evaluating ∇f at a point $\mathbf{x}$ gives the *gradient vector* of f at $\mathbf{x}$. The gradient vector is calculated as:

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix} .$$

Each of the components here is a *partial derivative*. To compute $\frac{\partial f(\mathbf{x})}{\partial x_1}$, take the derivative of f with respect to x_1, while treating all other variables as constants.

For example, let $f(\mathbf{x}) = x_1^2 + \sin(x_2) + x_2 x_3$. Then, the gradient of f evaluated at $\mathbf{x}$ is

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \frac{\partial f(\mathbf{x})}{\partial x_3} \end{pmatrix} = \begin{pmatrix} 2x_1 \\ \cos(x_2) + x_3 \\ x_2 \end{pmatrix} .$$

To compute the first entry of this vector, I treated x_2 and x_3 as constants, and took the derivative of f with respect to x_1. I repeated this process for each of the other two entries.