Vector Derivatives

Vectors

We can write partial derivatives in a more concise and elegant manner as follows

For a function $f : R^n \to R$ with value $f(x)$ for the vector $x \in R^n$, we put:

$$\nabla f = \frac{\partial f}{\partial x}$$

If f is an affine map then it is given by $f(x) = r + x \cdot c = r + c^T x = r + x^T c$ and we have:

$$\frac{\partial f}{\partial x} = \nabla f(x) = c$$

On the other hand, suppose $f : R^n \to R$ is quadratic function of the form $f(x) = x^T M x$ with $M \in R^{n \times n}$ and symmetric $M^T = M$. Then we have:

$$\frac{\partial f}{\partial x} = \nabla f(x) = 2Mx$$

Linear Regression

p = Number of features
n = Number of data Points

Create these matrices:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_p \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ 1 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

Calculate β

$$\beta = (X^T X)^{-1} X^T y$$