ANLY 590 HW 2.1 Jiali Chen

Part 1

1. $z = f(x, y) = ax + by + c$

$$\frac{\partial f}{\partial x} = a \qquad \frac{\partial f}{\partial y} = b$$

$$\nabla f(x, y) = \begin{bmatrix} a \\ b \end{bmatrix}$$

2 $z = f(x_1, x_2, \cdots, x_N)$

$$= \sum_{i=1}^{N} a_i (x_i - b_i) + S$$

$$= a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots + a_N x_N + d$$

$$\frac{\partial f}{\partial x_1} = a_1 \quad \frac{\partial f}{\partial x_2} = a_2 \quad \cdots \quad \frac{\partial f}{\partial x_N} = a_N$$

$$\nabla f(x_1, x_2, \cdots, x_N) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}$$

3 $z = f(x, y) = A(x - x_0)^2 + B(y - y_0)^2 + C$

$$f_x(x, y) = \left(\frac{\partial f(x, y)}{\partial x}\right)_y = 2A(x - x_0)$$

$$f_y(x, y) = \left(\frac{\partial f(x, y)}{\partial y}\right)_x = 2B(y - y_0)$$

4. $x^T = (3, 1, 4)$

$[1 \times 3]$

$$y^T = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$$

$[3 \times 1]$

$$B^T = \begin{pmatrix} 3 & 5 & 1 \\ 5 & 2 & 4 \end{pmatrix}$$

$[2 \times 3]$

$x \cdot x = 3 \times 3 + 1 \times 1 + 4 \times 4 = 9 + 1 + 16 = 26$

$x \cdot y^T = 3 \times 2 + 1 \times 5 + 4 \times 1 = 6 + 5 + 4 = 15$

$$x \times y = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \times (2 \ 5 \ 1) = \begin{pmatrix} 6 & 15 & 3 \\ 2 & 5 & 1 \\ 8 & 20 & 4 \end{pmatrix}$$

$[3 \times 3]$

$$y \times x = \begin{pmatrix} 2 & 5 & 1 \end{pmatrix} \times \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$$

$$= 2\times3 + 5\times1 + 1\times4 = (15)$$

$[1\times1]$

$$A \times x = \begin{pmatrix} 4\times3 + 5\times1 + 2\times4 \\ 3\times3 + 1\times1 + 5\times4 \\ 6\times3 + 4\times1 + 3\times4 \end{pmatrix} = \begin{pmatrix} 25 \\ 30 \\ 34 \end{pmatrix}$$

$[3\times1]$

$$A \times B = \begin{pmatrix} 4\times3+5\times5+2\times1 & 4\times5+5\times2+2\times4 \\ 3\times3+1\times5+5\times1 & 3\times5+1\times2+5\times4 \\ 6\times3+4\times5+3\times1 & 6\times5+4\times2+3\times4 \end{pmatrix} =$$

$$= \begin{pmatrix} 39 & 38 \\ 19 & 37 \\ 41 & 50 \end{pmatrix}$$

$[3\times2]$

B. reshape $(1,6) = (3, 5, 5, 2, 1, 4)$

$[1\times6]$

5. Loss Surface

$$L(p) = L(m,b) = \sum_{i=1}^{N} (\hat{y}_i - M(\hat{x}_i, m, b))^2$$

$$= \sum_{i=1}^{N} (y_i - m x_i - b)^2$$

Loss surface is minimized at

$$\frac{\partial L(m,b)}{\partial L(m)} = 0 \quad \text{and} \quad \frac{\partial L(m,b)}{\partial L(b)} = 0$$

$$\frac{\partial L(m,b)}{\partial L(b)} = -2 \sum_{i=1}^{N} (y_i - m x_i - b) = 0$$

$$\sum_{i=1}^{N} (y_i - m x_i - b) = 0$$

$$\sum_{i=1}^{N} y_i - m \sum_{i=1}^{N} x_i - Nb = 0$$

$$\frac{\sum_{i=1}^{N} y_i = m \sum_{i=1}^{N} x_i + Nb}{N}$$

$$\bar{y} = m \bar{x} + b$$

$$\boxed{b = \bar{y} - m \bar{x}}$$

$$\frac{\partial L(m,b)}{\partial L(m)} = 2\sum_{i=1}^{N}(y_i - mx_i - b)(-x_i) = 0$$

$$2\sum_{i=1}^{N}\left(-x_i y_i + m x_i^2 + b x_i\right) = 0$$

$$\frac{-2\sum_{i=1}^{N} x_i y_i + 2m\sum_{i=1}^{N} x_i^2 + 2b\sum_{i=1}^{N} x_i = 0}{2}$$

$$-\sum_{i=1}^{N} x_i y_i + m\sum_{i=1}^{N} x_i^2 + b\sum_{i=1}^{N} x_i = 0$$

$$\sum_{i=1}^{N} x_i y_i = m\sum_{i=1}^{N} x_i^2 + b\sum_{i=1}^{N} x_i$$

substitute b with $\overline{y} - m\overline{x}$

$$\sum_{i=1}^{N} x_i y_i = m\sum_{i=1}^{N} x_i^2 + \left(\overline{y} - m\overline{x}\right)\sum_{i=1}^{N} x_i$$

$$\sum_{i=1}^{N} x_i y_i = m\sum_{i=1}^{N} x_i^2 + \overline{y}\sum_{i=1}^{N} x_i - m\sum_{i=1}^{N} \overline{x} x_i$$

$$m\sum_{i=1}^{N} x_i^2 - m\sum_{i=1}^{N} \overline{x} x_i = \sum_{i=1}^{N} x_i y_i - \overline{y}\sum_{i=1}^{N} x_i$$

$$m\left(\sum_{i=1}^{N} x_i^2 - \sum_{i=1}^{N} \overline{x} x_i\right) = \sum_{i=1}^{N} x_i y_i - \overline{y}\sum_{i=1}^{N} x_i$$

$$m = \frac{\sum_{i=1}^{N} (x_i y_i - x_i \bar{y})}{\sum_{i=1}^{N} (x_i^2 - x_i \bar{x})}$$

since $\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$ $\quad \bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$

$$m = \frac{\sum_{i=1}^{N} x_i y_i - \frac{1}{N} \sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i^2 - \frac{1}{N} \sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} y_i}$$

$$m = \frac{\sum_{i=1}^{N} x_i y_i - \bar{x} \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i^2 - \frac{1}{N} \sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i}$$

$$m = \frac{\sum_{i=1}^{N} x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^{N} x_i^2 - n \bar{x}^2}$$

$$m = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\boxed{m = \frac{\mathrm{cov}(x, y)}{\mathrm{Var}(x)}}$$

$$\boxed{b = \bar{y} - \frac{\mathrm{cov}(x, y)}{\mathrm{Var}(x)} \bar{x}}$$

6. LLS Multi variable

Suppose a stack of N vectors $\{\vec{x_1}, \cdots \vec{x_N}\}$, each of dimension d and an scalar associated with each one $\{y_1, \cdots, y_n\}$

If we want a weight vector $\vec{w}$ such that

$$y_i \approx \vec{w}^T \vec{x}_i$$

we want to minimize the squared error using $\vec{w}$

$$\text{squared error} = \sum_{i=1}^{N} (y_i - \vec{x}_i \cdot \vec{w})^2$$

write it as vector equation:

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad X = \begin{bmatrix} - & \vec{x_1} & - \\ & \vdots & \\ - & \vec{x_N} & - \end{bmatrix}$$

$$\text{squared error} = \|Y - X\vec{w}\|^2$$

think of matrix X in terms of its d columns

$$X = \begin{bmatrix} | & & | \\ X_1 & \cdots & X_d \\ | & & | \end{bmatrix}$$

The optimal linear combination corresponds to dropping a line down from Y to $\{X_1 \ldots X_0\}$ at a right angle. In other words;

$$(Y - X\vec{w}) \cdot X_j = 0$$

for all columns $j=1$ to $j=d$, so

$$(Y - X\vec{w})^T X = \vec{0}$$

where $\vec{0}$ is d-component vector of zeros.

$$X^T(Y - X\vec{w}) = X^T Y - X^T X \vec{w} = 0$$

$$(X^T X)\vec{w} = X^T Y$$

$$\vec{w} = (X^T X)^{-1} X^T Y$$