HW-2.1 Assignment

Pedro Barbosa

PVB16

Part-1: Analytic assignment (10 points)

i.e. with pen and paper (show your work and all steps)

1.1 • Compute the gradient vector for a plane in 3D space (0.5 point)

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

Note:

$$\frac{d}{dx}x^2 = 2x$$

$$\frac{d}{dx}x^n = nx^{n-1}$$

1.2 • Compute the gradient vector for a hyperplane (0.5 point)

- Note: derivative of the sum is the sum of the derivatives

$$z = f(\mathbf{x}) = f(x_1, x_2, \ldots, x_N) = \sum_{i=1}^{N} a_i(x_i - b_i) + S = a_1x_1 + a_2x_2 + \cdots + a_Nx_N + d$$

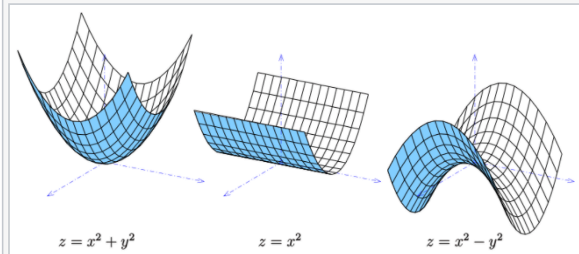
1.3 • Compute the partial derivative of the paraboloid function (1.5 point)

$$z = f(x, y) = A(x - x_o)^2 + B(y - y_o)^2 + C$$

$$f_x(x, y) = \left(\frac{\partial f(x, y)}{\partial x}\right)_y = ?$$

$$f_y(x, y) = \left(\frac{\partial f(x, y)}{\partial y}\right)_x = ?$$

Aside: Types of paraboloids



elliptic paraboloid, parabolic cylinder, hyperbolic paraboloid

1.4 • Given the following matrices and vectors (1.5 point)

$$\mathbf{x} = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 2 & 5 & 1 \end{pmatrix} \quad \mathbf{A} = \begin{pmatrix} 4 & 5 & 2 \\ 3 & 1 & 5 \\ 6 & 4 & 3 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 3 & 5 \\ 5 & 2 \\ 1 & 4 \end{pmatrix}$$

$$[3 \times 1] \quad [1 \times 3] \quad [3 \times 3] \quad [3 \times 2]$$

- compute the following quantities and specify the shape of the output, if an operation is not defined then just say "not defined"
  - (where dot specifies a dot product and x specifies a matrix product).

$$\mathbf{x}^T \quad \mathbf{y}^T \quad \mathbf{B}^T \quad \mathbf{x} \cdot \mathbf{x} \quad \mathbf{x} \cdot \mathbf{y}^T \quad \mathbf{x} \times \mathbf{y} \quad \mathbf{y} \times \mathbf{x} \quad \mathbf{A} \times \mathbf{x} \quad \mathbf{A} \times \mathbf{B} \quad \textbf{B.reshape(1,6)}$$

1.1 $\nabla z = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right)$; $z = ax + by + c$

$$\boxed{\nabla z = (a, b)}$$

1.2 $f(\bar{X}) = a_1x_1 + a_2x_2 + a_3x_3 \cdots + a_nx_n + d$

$$\nabla f(\bar{X}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \dots, \frac{\partial f}{\partial x_n}\right)$$

$$\boxed{\nabla f(\bar{X}) = (a_1, a_2, a_3, \dots, a_n)}$$

1.3 $z = A(x - x_0)^2 + B(y - y_0)^2 + C$

$$z = Ax^2 - 2Axx_0 + Ax_0^2 + By^2 - 2Byy_0 + By_0^2 + C$$

$$\boxed{\frac{\partial z}{\partial x} = 2Ax - 2Ax_0}$$

$$\boxed{\frac{\partial z}{\partial y} = 2By - 2By_0}$$

1.4

$$X = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \quad y = \begin{pmatrix} 2 & 5 & 1 \end{pmatrix} \quad A = \begin{pmatrix} 4 & 5 & 2 \\ 3 & 1 & 5 \\ 6 & 4 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 3 & 5 \\ 5 & 2 \\ 1 & 4 \end{pmatrix}$$

$$\boxed{X^T = \begin{pmatrix} 3 & 1 & 4 \end{pmatrix} \quad [1 \times 3]}$$

$$\boxed{y^T = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} \quad [3 \times 1]}$$

$$\boxed{B^t = \begin{pmatrix} 3 & 5 & 1 \\ 5 & 2 & 4 \end{pmatrix} \quad [2 \times 3]}$$

$$\boxed{X \circ X = \begin{pmatrix} 9 \\ 1 \\ 16 \end{pmatrix} \quad [3 \times 1]}$$

$$\boxed{X \circ y^t = \begin{pmatrix} 6 \\ 5 \\ 4 \end{pmatrix} \quad [3 \times 1]}$$

$$\boxed{X \times y = \begin{pmatrix} 6 & 15 & 3 \\ 2 & 5 & 1 \\ 8 & 20 & 4 \end{pmatrix} \quad [3 \times 3]}$$

$$\boxed{y \times X = \begin{pmatrix} 6 + 5 + 4 \end{pmatrix} = \begin{pmatrix} 15 \end{pmatrix} \quad [1 \times 1]}$$

$$\boxed{A \times X = \begin{pmatrix} 12 + 5 + 8 \\ 9 + 1 + 20 \\ 18 + 4 + 12 \end{pmatrix} = \begin{pmatrix} 25 \\ 30 \\ 34 \end{pmatrix} \quad [3 \times 1]}$$

$$\boxed{A \times B = \begin{pmatrix} 12 + 25 + 2, & 20 + 10 + 8 \\ 9 + 5 + 5, & 15 + 2 + 20 \\ 18 + 20 + 3, & 30 + 8 + 12 \end{pmatrix} = \begin{pmatrix} 39 & 38 \\ 19 & 37 \\ 41 & 50 \end{pmatrix} \quad [3 \times 2]}$$

$$\boxed{B.\mathrm{reshape}(1, 6) = \begin{pmatrix} 3 & 5 & 5 & 2 & 1 & 4 \end{pmatrix} \quad [1 \times 6]}$$

Part-1: Analytic assignment

Linear least squares (LLS): Single-variable (6 points)

1.5

- Use Calculus to analytically derive the expression for single variable linear regression fitting parameters using the sum of square error as the loss function (show your work)

Model: $y = M(x \,|\, \mathbf{p}) = mx + b$

$\mathbf{p} = (p_0, p_1) = (m, b)$

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

solution:

$$m = \frac{cov(x, y)}{var(x)} \qquad \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \text{var}(X) = \frac{1}{N}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$b = \bar{y} - \frac{cov(x, y)}{var(x)}\bar{x} \qquad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \qquad \text{cov}(X, Y) = \frac{1}{N}\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

Linear least squares (LLS): Multi-variable (EXTRA CREDIT) (+3 points)

- Use matrix calculus to analytically derive the expression for **two variable** linear regression fitting parameters using the sum of square error as the loss function
  - Show your work using matrix notation
  - From your solution infer the generalized solution for an arbitrary number of variables

solution: $\vec{w} = (X^\top X)^{-1} X^\top Y.$

1.5 $L = \sum_{i=1}^{N} \left(\hat{y}_i - (m\hat{X}_i + b)\right)^2$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{N} -2(\hat{y}_i - m\hat{X}_i - b) = 0 \quad \div(-2)$$

$$\sum_{i=1}^{N} (\hat{y}_i - m\hat{X}_i - b) = 0 \ ; \ \sum_{i=1}^{N} (-b) = -Nb$$

$$\sum_{i=1}^{N} \hat{Y}_i - m\sum_{i=1}^{N} \hat{X}_i - Nb = 0 \quad +(Nb) \quad \div(N)$$

$$\frac{\sum_{i=1}^{N} \hat{Y}_i - m\sum_{i=1}^{N} \hat{X}_i}{N} = b \ ; \ \frac{\sum_{i=1}^{N} \hat{Y}_i}{N} = \text{mean}(Y) = \overline{Y}$$

(same for X)

$$\boxed{b = \overline{Y} - m\overline{X}} \quad (1)$$

$$\frac{\partial L}{\partial m} = \sum_{i=1}^{N} -2\hat{X}_i(Y_i - b - mX_i) = 0 \quad \div(-2)$$

$$\sum_{i=1}^{N} (\hat{X}_i\hat{Y}_i - b\hat{X}_i - m\hat{X}_i^2) = 0 \text{, substituting (1)}$$

$$\sum_{i=1}^{N} (\hat{X}_i\hat{Y}_i - (\overline{Y} - m\overline{X})\hat{X}_i - m\hat{X}_i^2) = 0 \text{, rearranging:}$$

$$\sum_{i=1}^{N} (\hat{X}_i\hat{Y}_i - \overline{Y}\hat{X}_i) - m\underbrace{\sum_{i=1}^{N} (\hat{X}_i^2 - \overline{X}X_i)}_{} = 0$$

$$m\sum_{i=1}^{N}\left(\hat{x}_i^2-\bar{x}\,\hat{x}_i\right)=\sum_{i=1}^{N}\left(\hat{x}_i\hat{y}_i-\bar{y}\,\hat{x}_i\right)$$

$$\boxed{m=\frac{\sum_{i=1}^{N}\left(\hat{x}_i\hat{y}_i-\bar{y}\,\hat{x}_i\right)}{\sum_{i=1}^{N}\left(\hat{x}_i^2-\bar{x}\,\hat{x}_i\right)}=\frac{\operatorname{cov}(x,y)}{\operatorname{var}(x)}} \quad (2)$$