NLP: Homework 1: Background Review + Building a classifier

Raef Youssef

2024/01/26

1 Refreshing the Rows and Columns: Linear Algebra Review

1.1 Basic Operations

- Givens

$$\alpha = 2, x = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, y = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}, z = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, A = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

1. $$x * y = 0 + 4 + 10 = 14$$
2. $$x * z = 0 + 2 - 2 = 0$$
3. $$\alpha(x + y) = \begin{bmatrix} 6 \\ 10 \\ 14 \end{bmatrix}$$
4. $$||x|| = \sqrt{5}$$
5. $$x^T = \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}$$
6. $$Ax = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 3 & 1 \\ 2 & 2 & 6 \end{bmatrix}$$
7. $$x^T Ax = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 2 & 12 \end{bmatrix}$$

1.2 Matrix Algebra Rules

1. True
2. True
3. False
4. True
5. False
6. True
7. True
8. True
9. True

1.3 Matrix operations

- Givens

$$B = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

- Is B invertible? If so find B^{-1}

$$det(B) = -(-1)(-1) + (1)(2-1) = -1 + 1 = 0$$

  B is not invertible

- Is B diagonalizable? If so, find its diagonalization
  - Eigenvalues

$$det(B - \lambda I) = 0 \longrightarrow \lambda_1 = 0, \lambda_2 = 2, \lambda_3 = 2$$

  - Eigenvectors

$$(B - 0I)v = 0 \longrightarrow v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$(B - 2I)v = 0 \longrightarrow v_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, v_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

  - Calculate P

$$P = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

– Calculate D

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

– Diagonalization

$$B = PDP^{-1}$$

2 Taking Chances: Probability Review

2.1 Basic Probability

1. Tickets should be $5

$$P(win) = 12/36 = 1/3$$
$$P(loss) = 1 - P(win) = 2/3$$
$$EV = 15(1/3) + 0(2/3) = 5$$

2. P(B) = .95 - .4 = .55

3.

$$P(A \cup B) = P(A) + P(B) - P(A)P(B)$$
$$.95 = .4 + .6P(B) \longrightarrow P(B) = .9167$$

2.2 Expectations and Variance

1.

$$P(H) = .5(.3 + .9) = .6$$
$$E[X] = (3)(.6) = 1.8$$

2.

$$Var(X) = 3(.6)(.4) = .72$$

2.3 A Variance Paradox?

1. There's no contradiction. When you sum i.i.d. random variables, you can sum their variance. However, when you add a random variable to itself, it's not i.i.d, instead you scale its values by $2x$. Therefore $Var(X + X) = Var(2X) = 4Var(X)$

3 Calculus Review

3.1 One-variable derivatives

- $$f^{'}(x) = 6x - 2$$
- $$f^{'}(x) = 1 - 2x$$
- $$f^{'}(x) = 1 - \frac{\exp(-x)}{1 + \exp(-x)} = 1 - (1 - p(x)) = p(x)$$

3.2 Multi-variable derivative

- $$\Delta f(x) = [x_1^2, exp(x_2)]$$
- $$\Delta f(x) = exp(x_1, x_2x_3), x_3exp(x_1, x_2x_3), x_2exp(x_1, x_2x_3)]$$
- $$\Delta f(x) = a$$
- $$\Delta f(x) = 2Ax = \begin{bmatrix} 4x_1 - 2x_2 \\ -2x_1 + 2x_2 \end{bmatrix}$$
- $$\Delta f(x) = x$$

4 Algorithms and Data Structures Review

1. $$O(nlog(n))$$
2. $$O(n)$$
3. $$O(log(n))$$
4. $$O(1)$$
5. $$O(nd)$$

6.

$$O(d^2)$$

7.

$$O(mnd)$$