Q5

April 30, 2024

0.1 a)

The PCA is conduct by obtaining new parameters z by mapping them with the old paramters x in a linear relation.

The correlation of z and x is linked by a loading weight w, saying how much weighting of each old parameters is considred by the new parameter z.

$$z_1 = w_1^T x$$

To find the new z, we try to maximize the variance of the new z, such that it can capture the most information from the old x into the new z, hence reducing the dimension of the paramter space.

$$\begin{aligned} Var(z_1) &= Var(w_1^T x) \\ &= E((w_1^T x - w_1^T \mu)^2) \\ &= w_1^T E((x-\mu)(x-\mu)^T) w_1 \\ &= w_1^T \Sigma w_1 \end{aligned}$$

Since we try the maximize the loading weight, it has to be set on a constraint to prevent the weights bloating into excessive value. We try to limit the weights by noramlizing its distance into 1:

$$||w_1|| = \sum_n w_{n1}^2 = 1$$

This singly constrainted linear programming can be solved by differentiating the Lagrangian equation of the model with respect to the w_1

$$\mathcal{L}(w_1) = Var(z_1) - \alpha(||w_1|| - 1)$$

$$= Var(w_1^T x) - \alpha(||w_1|| - 1)$$

$$= w_1^T \Sigma w_1 - \alpha(w_1^T w_1 - 1)$$

$$\frac{\partial \mathcal{L}(w_1)}{\partial w_1} = \Sigma w_1 - \alpha w_1 = 0$$

$$(\Sigma - \alpha I)w_1 = 0$$

This implies α, as a scalar variable, should be the eigenvalue and w_1 should be a eigenvector for Σ, a nxn matrix where n is the dimension of x, to be reduced into the formula.

1 b)

- PCA can lower the dimension of the variable vector, hence saving the calculation times and resources for the model prediction.
- It can reduce the variable counts, hence identify the most critical variables out of the provided.
- By looking at the loading vector of the variables to the principal, one can identify the collinearity of the variables if the loading vectors are similar

2 c)

$$X = \begin{bmatrix} x & y & z \end{bmatrix} = \begin{bmatrix} -1 & -1 & 2 \\ -2 & 1 & 1 \end{bmatrix}$$

$$\mu_x = \frac{\sum_n x_n}{n} = \frac{-1-2}{2} = -1.5$$

$$\mu_y = \frac{\sum_n x_n}{n} = \frac{-1+1}{2} = 0$$

$$\mu_z = \frac{\sum_n x_n}{n} = \frac{2+1}{2} = 1.5$$

Calculate the Covariance

$$\sum = \frac{1}{N} \sum_n (x_n - \mu)(x_n - \mu)^T$$

$$= \begin{bmatrix} var(x) & cov(x,y) & cov(x,z) \\ cov(y,x) & var(y) & cov(y,z) \\ cov(z,x) & cov(z,y) & var(z) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{(-1+1.5)^2+(-2+1.5)^2}{2} & \frac{(-1+1.5)(-1-0)+(-2+1.5)(1-0)}{2} & \frac{(-1+1.5)(2-1.5)+(-2+1.5)(1-1.5)}{2} \\ \frac{(-1-0)(-1+1.5)+(1-0)(-2+1.5)}{2} & \frac{(-1-0)^2+(1-0)^2}{2} & \frac{(-1-0)(2-1.5)+(1-0)(1-1.5)}{2} \\ \frac{(2-1.5)(-1+1.5)+(1-1.5)(-2+1.5)}{2} & \frac{(2-1.5)(-1-0)+(1-1.5)(1-0)}{2} & \frac{(2-1.5)^2+(1-1.5)^2}{2} \end{bmatrix}$$

$$= \begin{bmatrix} 0.25 & -0.5 & 0.25 \\ -0.5 & 1 & -0.5 \\ 0.25 & -0.5 & 0.25 \end{bmatrix}$$

The derivitive of the Lagrangian will yield the follwoing equation with the eigenvalue to be solved:

$$\Sigma w = \alpha w$$

Rewrite the eigen equation:

$$(\Sigma - \alpha I)w = 0$$

$$\begin{vmatrix} 0.25-\alpha & -0.5 & 0.25 \\ -0.5 & 1-\alpha & -0.5 \\ 0.25 & -0.5 & 0.25-\alpha \end{vmatrix} = 0$$

$$= (0.25-\alpha)\begin{vmatrix} 1-\alpha & -0.5 \\ -0.5 & 0.25-\alpha \end{vmatrix} - (-0.5)\begin{vmatrix} -0.5 & -0.5 \\ 0.25 & 0.25-\alpha \end{vmatrix} + 0.25\begin{vmatrix} -0.5 & 1-\alpha \\ 0.25 & -0.5 \end{vmatrix}$$

$$(0.25-\alpha)(\alpha^2 - 1.25\alpha + 0.25 - 0.25) + 0.5(-0.125 + 0.5\alpha + 0.125) + 0.25(0.25 - 0.25 + 0.25\alpha)$$

$$= -\alpha^3 + 1.5\alpha^2 - 0.3125\alpha + 0.25\alpha + 0.0625\alpha$$

$$= -\alpha^3 + 1.5\alpha^2$$

$$= -\alpha^2(\alpha - 1.5)$$

$\alpha = 1.5$ or $\alpha = 0$

Choosing the largest eigenvalue $\alpha = 1.5$ for maximizing the variance

Substitube back to the eigen equation

$$(\Sigma - \alpha I)w = (\begin{bmatrix} 0.25 & -0.5 & 0.25 \\ -0.5 & 1 & -0.5 \\ 0.25 & -0.5 & 0.25 \end{bmatrix} - \begin{bmatrix} 1.5 & 0 & 0 \\ 0 & 1.5 & 0 \\ 0 & 0 & 1.5 \end{bmatrix})w = \begin{bmatrix} -1.25w_1 - 0.5w_2 + 0.25w_3 \\ -0.5w_1 - 0.5w_2 + -0.5w_3 \\ 0.25w_1 - 0.5w_2 - 1.25w_3 \end{bmatrix} = 0$$

Assume $w_1 = k$, from the eigenequations

$$w_2 = -2k,\ w_3 = k$$

$$w = k\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Normalize the vector s.t. $||w|| = 1$

$$||w|| = k\sqrt{1^2 + (-2)^2 + 1^2} = 1$$

$$k = \frac{1}{\sqrt{6}}$$

$$w = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$$