Exercises for Chapter 4, Part 1

4.1 Let $X = \{(-2, 1), (-1, 2), (0, 0), (1, -2), (2, -1)\}$.

a) Sketch a scatter diagram of this data set and the eigenvectors of its covariance matrix.

b) Compute the result of one-dimensional principal component projection.

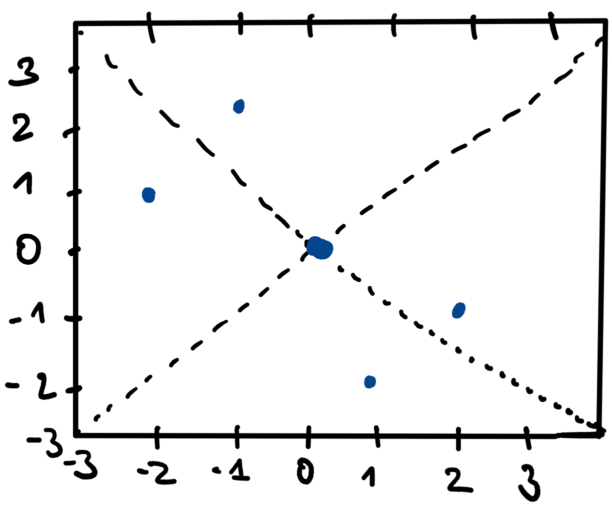
c) Compute the average quadratic projection error.

d) Sketch a Shepard diagram of this projection.

$X = \{(-2,1), (-1,2), (0,0), (1,-2), (2,-1)\}$

$\bar{x} = (0,0)$

a)



Because of symmetry, the mean of the data is at the origin. The symmetry axis along which the variance is maximum is the +45° line. Any derivation from this line will reduce the variance, so the eigenvector with the maximum eigenvalue is in direction $(1,-1)$, and normalization yields $\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$. Eigenvectors are pairwise orthogonal. Hence, if the first eigenvector is along the -45° line, then the second eigenvector is along the +45° line, which is in direction $(1,1)$, and normalization yields $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$. The direction of any of the eigenvectors may be reverted.

b)

$(-2,1)^T \cdot \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = -\frac{3}{2}\sqrt{2}$ $\quad$ $(1,-2)^T \cdot \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = \frac{3}{2}\sqrt{2}$

$(-1,2)^T \cdot \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = -\frac{3}{2}\sqrt{2}$ $\quad$ $(2,-1)^T \cdot \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = \frac{3}{2}\sqrt{2}$

$(0,0)^T \cdot \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = 0$ $\quad \Rightarrow Y = \left\{-\frac{3}{2}\sqrt{2}, -\frac{3}{2}\sqrt{2}, 0, \frac{3}{2}\sqrt{2}, \frac{3}{2}\sqrt{2}\right\}$

c)

Reverse projection of Y yields

$-\frac{3}{2}\sqrt{2} \cdot \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = \left(-\frac{3}{2}, \frac{3}{2}\right)$ $\quad$ $\frac{3}{2}\sqrt{2} \cdot \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = \left(\frac{3}{2}, -\frac{3}{2}\right)$

$0 \cdot \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = (0,0)$ $\quad \Rightarrow X' = \left\{\left(-\frac{3}{2}, \frac{3}{2}\right), \left(\frac{3}{2}, \frac{3}{2}\right), (0,0), \left(\frac{3}{2}, -\frac{3}{2}\right), \left(\frac{3}{2}, -\frac{3}{2}\right)\right\}$

Square distances:

$\|(-2,1) - (-\frac{3}{2}, \frac{3}{2})\|^2 = 1/2$ $\quad$ $\|(1,-2) - (\frac{3}{2}, -\frac{3}{2})\|^2 = 1/2$

$\|(-1,2) - (-\frac{3}{2}, \frac{3}{2})\|^2 = 1/2$ $\quad$ $\|(2,-1) - (\frac{3}{2}, -\frac{3}{2})\|^2 = 1/2$

$\|(0,0) - (0,0)\|^2 = 0$ $\quad$ average distance $e = \frac{1}{5}\left(\frac{1}{2} + \frac{1}{2} + 0 + \frac{1}{2} + \frac{1}{2}\right) = \frac{2}{5}$

d) Shepard diagram

The pairwise distance matrices of X and Y are

distance between x_1 and x_2

$$D^X = \begin{pmatrix} 0 & \sqrt{2} & \sqrt{5} & 3\sqrt{2} & 2\sqrt{5} \\ \sqrt{2} & 0 & \sqrt{5} & 2\sqrt{5} & 3\sqrt{2} \\ \sqrt{5} & \sqrt{5} & 0 & \sqrt{5} & \sqrt{5} \\ 3\sqrt{2} & 2\sqrt{5} & \sqrt{5} & 0 & \sqrt{2} \\ 2\sqrt{5} & 3\sqrt{2} & \sqrt{5} & \sqrt{2} & 0 \end{pmatrix}$$

$$D^Y = \begin{pmatrix} 0 & 0 & \frac{3}{2}\sqrt{2} & 3\sqrt{2} & 3\sqrt{2} \\ 0 & 0 & \frac{3}{2}\sqrt{2} & 3\sqrt{2} & 3\sqrt{2} \\ \frac{3}{2}\sqrt{2} & \frac{3}{2}\sqrt{2} & 0 & \frac{3}{2}\sqrt{2} & \frac{3}{2}\sqrt{2} \\ 3\sqrt{2} & 3\sqrt{2} & \frac{3}{2}\sqrt{2} & 0 & 0 \\ 3\sqrt{2} & 3\sqrt{2} & \frac{3}{2}\sqrt{2} & 0 & 0 \end{pmatrix}$$

This yields a Shepard diagram with four points at $(\sqrt{2}, 0)$ (two-fold), $(\sqrt{5}, \frac{3}{2}\sqrt{2})$ (fourd-fold) $(3\sqrt{2}, 3\sqrt{2})$ (two-fold) $(2\sqrt{5}, 3\sqrt{2})$ (two-fold