Exercise 2 - Principal Component Analysis

Giulio Valente

1 Problem 1

1.1 1

1.1.1 a

$$\bar{\mathbf{X}} = \mathbf{X} - \mathbf{M}$$

1.1.2 b

$$\frac{1}{N-1}\bar{\mathbf{X}}\bar{\mathbf{X}}^T$$

Definition of covariance

$$cov(u,v) = E[(u - E[v])(v - E[v])]$$

Where the expectation has some distribution

$$\frac{1}{n-1}(u-\mu_1)(v-\mu_1)$$

1.1.3 c

$$\mathbf{\Sigma} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$$

U is the eigenvector matrix $\in R^{dxd}$
Σ is the diagonalized eigenvalue matrix

1.1.4 d

$$\bar{\mathbf{Z}} = \mathbf{U}_K^T\bar{\mathbf{X}}$$

$\mathbf{U}_K \in R^{DxK}$ just takes the first K eigenvectors $\mathbf{Z}_K \in R^{KxN}$ Every measurements is just a projection that use the k eigenvector.

1.1.5 e

$$\tilde{\mathbf{X}_K} = \mathbf{U}_K(\mathbf{U}_K)^T\mathbf{Z} + \mu(???)$$
$$\tilde{\mathbf{X}_K} \in R^{DxN}$$

1.2 2

1.2.1 a

The number of dimensionality appears to be high. For every dimensionality we have different eigenvalues.

1.2.2 b

No

1.2.3 c

None

1.3 3

1. B
2. E
3. C No correlation between the two dimensions.

1.4 4

1.4.1 a

TODO

1.4.2 b

$$\mathbf{\Sigma}_z = \frac{1}{N+1}(z-\mu_z)(z-\mu_z)^T =$$
$$= (\mathbf{A}^T\mathbf{A} - \mathbf{A}\mu_x)(\mathbf{A}^T\mathbf{X} - \mathbf{A}^T\mu_x)^T =$$
$$= \mathbf{A}^T(x-\mu_x)(\mathbf{A}^T(x-\mu_x))^T =$$
$$= \frac{1}{N+1}\mathbf{A}^T(x-\mu_x)(x-\mu_x)^T)A = \mathbf{A}^T\mathbf{\Sigma}_x\mathbf{A}$$

1.4.3 c

$$\mathbf{\Sigma}_Z = \mathbf{U}^T(\mathbf{U}\Lambda\mathbf{U}^T)U = \mathbf{\Lambda}$$