Paul Johnston
CS 478
2/13/19

# PCA Hw

Original Data

| | x | y |
|---|---|---|
| P1 | .2 | -.3 |
| P2 | -1.1 | 2 |
| P3 | 1 | -2.2 |
| P4 | .5 | -1 |
| P5 | -.6 | 1 |
| mean | 0 | -.1 |

Terms

| | | |
|---|---|---|
| m | 5 | number of instances in data set |
| n | 2 | Number of input features |
| P | 1 | Final number of principal components chosen |

**Step 1** center

Centered Data

| | x' | y' |
|---|---|---|
| P1 | .2 | -.2 |
| P2 | -1.1 | 2.1 |
| P3 | 1 | -2.1 |
| P4 | .5 | -.9 |
| P5 | -.6 | 1.1 |

**Step 2**: Calculate covariance matrix

$$cov(x,y) = \frac{(.2 \cdot -.2)+(-1.1 \cdot 1.9)+(1 \cdot -2.1)+(.5 \cdot -.9)+(-.6 \cdot .9)}{4}$$

$$= -1.305$$

$$cov(x,x) = \frac{.2^2 + -1.1^2 + 1^2 + .5^2 + -.6^2}{4}$$

$$= 0.715$$

$$cov(y,y) = \frac{-.2^2 + 1.9^2 + -2.1^2 + -.9^2 + .9^2}{4}$$

$$= 2.42$$

$$cov = \begin{bmatrix} 0.715 & -1.39 \\ -1.39 & 2.72 \end{bmatrix}$$

**Step 3** calculate unit eigenvectors

```
import numpy as np
cov = np.array([[0.715, -1.39 ],
                [-1.39 , 2.72]])
w, v = np.linalg.eig(cov)
```

$$\text{eigenvalues} = \begin{bmatrix} 0.00370122 \\ 3.43129878 \end{bmatrix}$$

$$\text{eigenvectors} = \begin{bmatrix} -0.89021285 & 0.45554483 \\ -0.45554483 & -0.89021285 \end{bmatrix}$$

**Step 4** Keep P eigenvectors

$$A = \begin{bmatrix} 0.45554483 & -0.89021285 \end{bmatrix}$$

**Step 5** Transform

$$B = \begin{bmatrix} .2 & -1.1 & 1 & .5 & -.6 \\ -.2 & 2.1 & -2.1 & -.9 & 1.1 \end{bmatrix}$$

$$T = A \times B$$

$$T^T = \begin{matrix} 0.269151536 \\ -2.370546298 \\ 2.324991815 \\ 1.02896398 \\ -1.252561033 \end{matrix}$$

$$\frac{3.43\ldots}{3.43\ldots + .0037} = 99.89\ \%$$

of total info is contained