

Lecture Three

March 3, 2016

1 Principal Component

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (\mu^T (\vec{x}_i - \vec{\mu})(\vec{x}_i - \vec{\mu})^T) \vec{\mu} \quad (1)$$

1.1 eigenvalues & eigenvectors

$$A\vec{x} = \lambda\vec{x} \quad (2)$$

The eigen vectors are the vectors that when scales do not effect the direction of the vectors. The eigen vectors will tell you the rank of your matrix and help in discerning that your covariance matrix will be singular and not invertible. With this you are able to reduce the number of features that you consider.

If you have more features than you would like you should consider removing the features that have small variance. The small variance could be noise in the data and considering it could degrade the classification. The larger the eigen value the more variance that is associated with the data. Small eigen values are not associated with good data.

If the data is centered than you can write all of the features as a linear combination of the eigenvectors.

$$\vec{x} = \alpha_1 \vec{e}_1 + \alpha_2 \vec{e}_2 + \dots \alpha_n \vec{e}_n \quad (3)$$

$$\vec{u}^T \Sigma \vec{u} = \vec{u}(\alpha_1 \lambda_1 \vec{e}_1 + \alpha_2 \lambda_2 \vec{e}_2 + \dots + \alpha_n \lambda_n \vec{e}_n) \quad (4)$$

$$= \alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2 + \dots + \alpha_n^2 \lambda_n \quad (5)$$

the inverse matrix has the same eigen vectors of the original matrix.

2 Fisher Discriminant

Take a problem in a R^n space and transform it into a R^1 space. Then you can fit two gaussians that represent the classifications.

Criteria:

1. max separation of means

2. smallest projected variance

$$S(\vec{u}) = \frac{||\vec{a}\vec{\mu}_1 - \vec{a}\vec{\mu}_2||^2}{\vec{a}^T \Sigma_1 \vec{a} + \vec{a}^T \Sigma_2 \vec{a}} \quad (6)$$

Maximize equation to find a good compromise between the smallest projected variance and the max separation of means.

$$\vec{a} = c * (\Sigma_1 + \Sigma_2)^{-1}(\vec{\mu}_1 - \vec{\mu}_2) \quad (7)$$

Fisher Classifier:

1. compute: $\Sigma_1, \Sigma_2 | \mu_1, \mu_2$
2. compute: \vec{a}
3. One dimensional problem
4. apply your favorite classifier (gaussian)

For the homework apply the Gaussian classifier to the linearized data.

3 Next Lecture

Expectation Maximization algorithm.