总结报告8

（2019.10.9——2019.10.10）

**一、学习内容：机器学习中的数学：PCA WEEK4——PCA Algorithm；**

**PCA: an algorithm for linear dimensionality reduction**

**A key idea: to use orthogonal projections to find lower dimensional representations of data that remain as much as possible**

1. **Finding the coordinates of projected data**













 span the principal subspace



1. **Optimal projection parameters**

**Assume**

** **

****the optimal coordinates****

**3．Reformulation of the objective**





****





 trace：迹



J：loss function 数据投影到忽略子空间上的方差

1. **Finding the basis vectors that span the principal subspace**

**KEY IDEA:** Minimizing the average squared reconstruction error

minimizing the projection of the variance of the data by projecting it onto subspace that we ignore in PCA.

 S:data covariance matrix

考虑2维 



**Lagrange multiplier:**



**Notice:** the eigen vectors of the covariance matrix are already orthogonal to each other(symmetric)

**General Case:**



**Notice:** The orthonormal basis of principal subspace as the eigenvectors of the data covariance matrix that are associated with the largest eigenvalues.

**PCA algorithm**

**5．steps of PCA**

Comments：

1. E[X]=0 is not necessary.
2. Recommend: unit-free



B: the matrix that contains the eigenvectors that belong to the largest eigenvalues as columns.

: the coordinates of the projection with respect to the basis of the principal subspace

**PCA STEPS:**

1. subtract **the mean** from the data and send it at zero to avoid numerical problems.
2. divide by the **standard deviation** to make the **data unit-free**.
3. compute the **eigenvalues and eigen vectors** of the **data covariance matrix**.

④ **project any data point onto the principal subspace** that is spanned by the eigenvectors that belong to the largest eigenvalues.

1. **PCA in high dimensions**

****



**** : eigenvectors of **S**

**Remark:** we need to normalize the eigenvectors. (ONB properties)

The data is substantially(实质上，充分地)bigger than the number of data points.

1. **Other interpretations(解释) of PCA**
   1. minimising the squared reconstruction error
   2. minimising the autoencoder loss
   3. mazimising the mutual information
   4. mazimising the variance of the projected data
   5. mazimising the likelihood in a latent variable model（latent ：adj. 潜在的）