Dimensional reduction

Why?

* Easily visualize < 3 Dim
* Reduce computation cost
* Reduce overfitting
* *High - level embedding*, compare

How?

* Maximize variance – High distance between data point

Khong mat tinh tong quat

Let

Suppose:

We now maximize

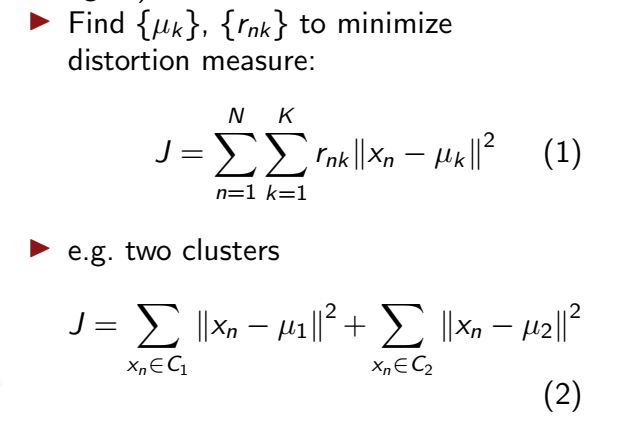
Lagrange function:

Elbow PCA:

**K-mean**

Hard-cluster: Each datapoint belong to 1 class

Soft cluster: Each datapoint can belong to many classes



Fix → find .

Gaussian Mixture Model

Underfitting: accuracy thấp

Overfitting: accuracy trên train và test lệch

Stop overfit

Iteration

Learning rate

L1, L2

Weight error model

**Ensemble method**

Abstract,

Each single model have their own weakness (high bias/variance), so we should combine them.

How: Chose base model → train in different ways to reduce bias/variance

* (Variance) Bagging: same type models are trained in different subsamples from origin train set( randomly) parallelly, then take the average outcome
* (Main Bias) Boosting: same type models train consecutively, after learn from error of before model → the output is sum of
* (Main bias) Stacking: varied type models + meta model (supervised model)

Sample data = Bootstrapping: Random order of origin data

<https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/>

model selection + tuning params: <https://towardsdatascience.com/ensemble-model-selection-evaluation-in-machine-learning-by-optimalflow-9e5126308f12>

bootstrapping: <https://towardsdatascience.com/an-introduction-to-the-bootstrap-method-58bcb51b4d60>

ts: <https://towardsdatascience.com/time-series-in-python-exponential-smoothing-and-arima-processes-2c67f2a52788>

<https://phamdinhkhanh.github.io/2019/12/12/ARIMAmodel.html>

