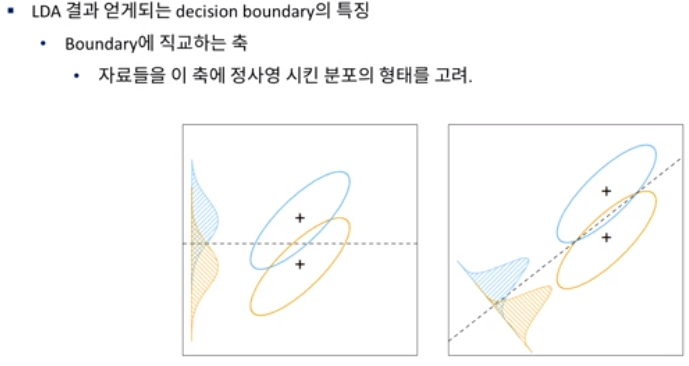
**<LDA (linear discriminant analysis)>**

가정

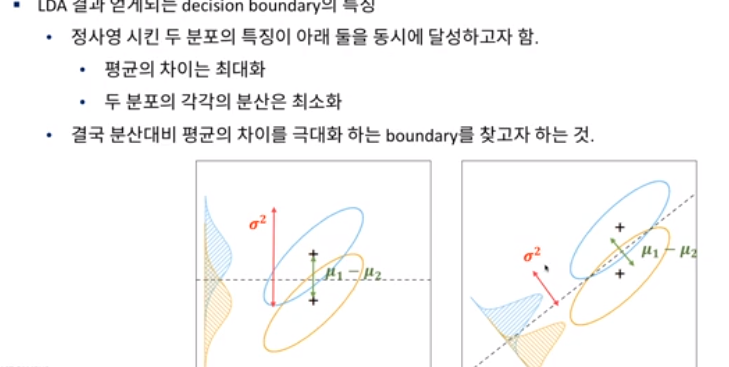
* + 각 집단은 정규분포 형태의 확률분포를 가진다
  + 각 집단은 비슷한 형태의 공분산 구조를 가진다.



각 집단의 분포를 표현하고 집단의 평균값들의 중심값을 boundary로 잡는다

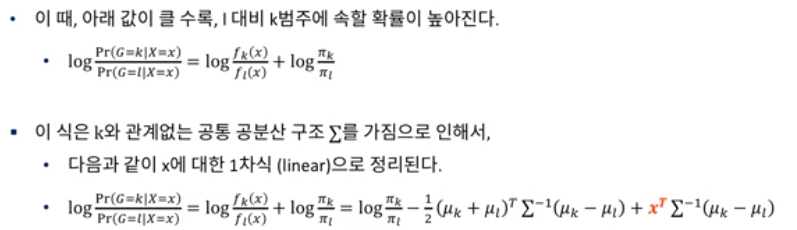
자료가 같음에도 정사영을 내린 축이 달라지면 boundary가 바뀐다.

* 어떤 축에 정사영을 내려하는가? 평균의 차이를 가장 크게 하는 축으로 내린다



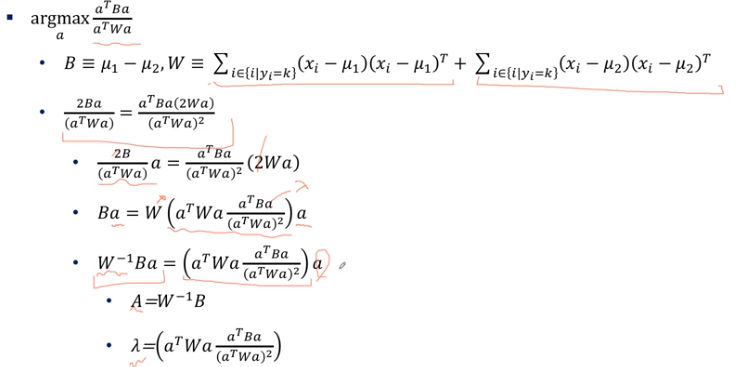
타원의 단축방향으로 정사영을 내린다.

정규분포를 통한 이변량 정규분포. 식은 언급만 **<LDA 수학적 개념이해>강의 참고**



공통 공분산 구조를 가지지 않으면 소거가 되지않는다 ( QDA )

**<사영>**

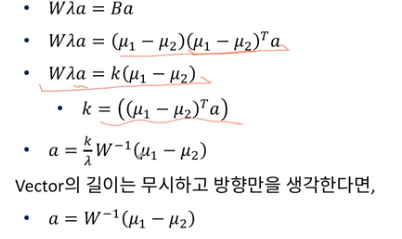


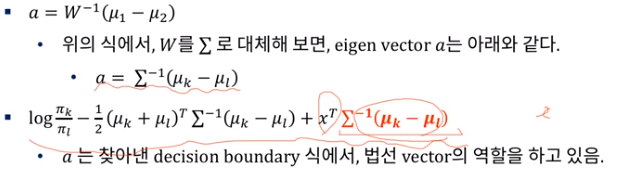
최종 식에서 eigen vector와 eigen value로 생각할수 있음

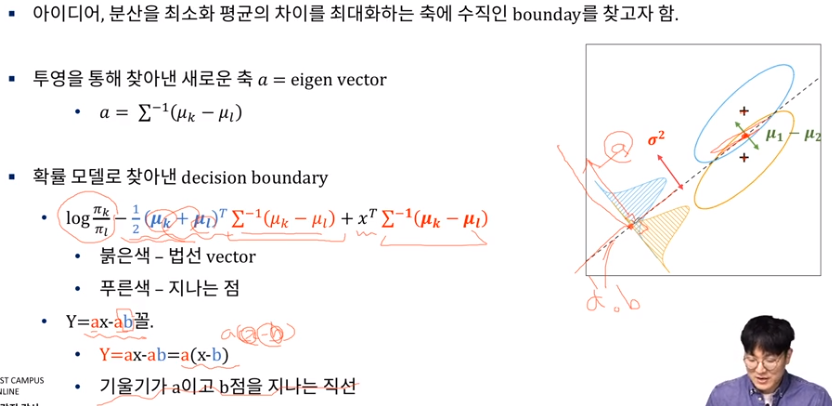
a라는 벡터가 양변에 존재하고 왼쪽의 WB는 매트릭스의 형태, 오른쪽의 식은 스칼라로 eigen value

따라서 a는 eigen vector

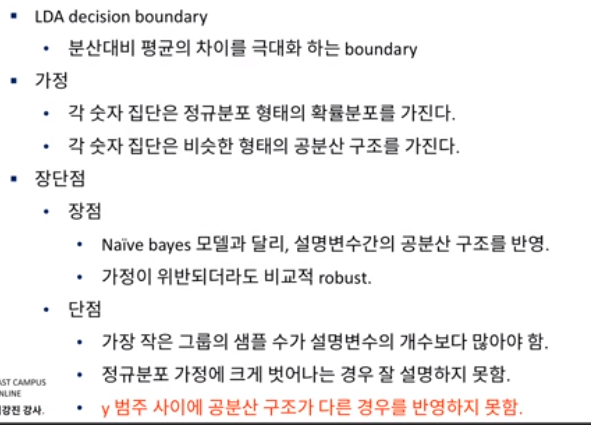
* LDA의 개념 : 확률의 비율에 로그를 씌운 뒤 0보다 크면 분자의 범주, 0보다 작으면 분모의 범주. 분산과 평균의 비율을 최대화하는( 분산은 최소화, 평균의 차는 최대 )





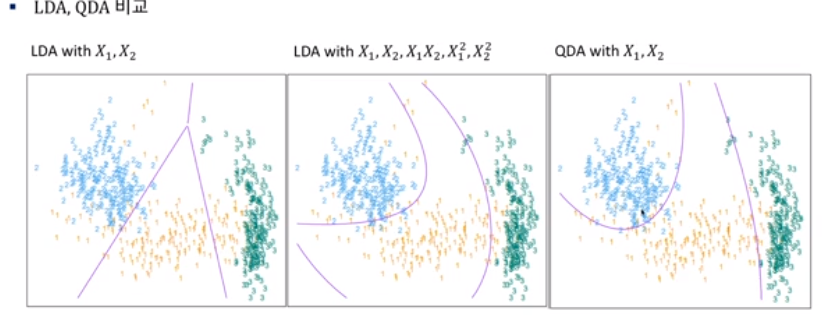


**<LDA 정리>**



QDA란

K와 관계없는 공통 공분산 구조 (시그마)에 대한 가정을 버린 것 : y의 범주별로 서로 다른 공분산 구조를 가진 경우에 활용 가능 식에서 사라지는 (시그마)부분이 없음



2번째 분류는 공분산구조가 같다고 가정하고 모델을 측정한 것.

3번째는 QDA로 각각 다른 공분산 구조여도 가능함. 대신 샘플이 충분해야한다.

