SVM

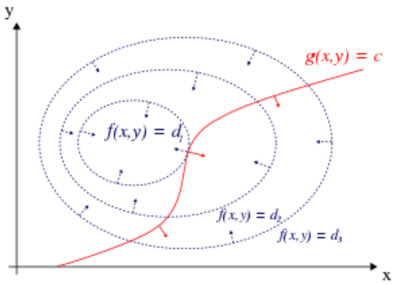
Lagrangian approach

: **연립방정식 제한조건**, 즉 equality constraint (등식 제한조건)이 있는 최적화 문제를 푸는 방법. 라그랑지안 승수 (multiplier) 항을 더하여 제약된 문제를 제약이 없는 문제로 바꾸는 방법

$$\mathcal{L}(x,\lambda) = f(x) - \lambda g(x)$$

$\mathcal{L}$ = Lagrangian
λ = Lagrange multiplier
$g(x)$ = equality constraint
$f(x)$ = function
x = integer

- 물체의 에너지로부터 운동 방정식 유도. 유체 역학에서 액체가 움직이는 양상 모델링할 때 쓰임. “운동 에너지와 위치 에너지의 차”
- 원래의 목적함수 $f(x)$를 고려하는대신, 제한조건 $g(x)$에 λ를 곱한 것을 더한 함수 $f(x) + \lambda g(x)$를 정의해서 최적화



KKT (Karush-Kuhn-Tucker) Condition

: **연립부등식 제한조건** (inequality constraint)이 있는 최적화 문제에서 만족하도록 하는 추가적인 수식 조건

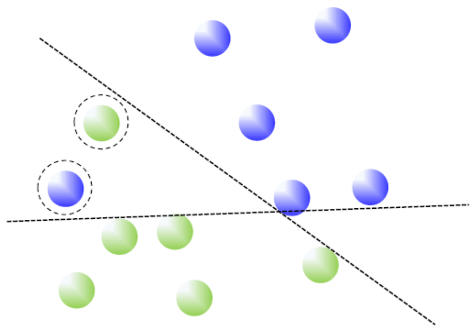
Duality (쌍대성)

: 최적화 문제를 **원초문제 (primal problem)**과 **쌍대문제 (dual problem)**의 두가지 관점에서 볼 수 있다는 원칙

- 라그랑지안 접근법으로 다양한 유형의 primal problem을 dual problem으로 바꿔서 풀 수 있음

SVM

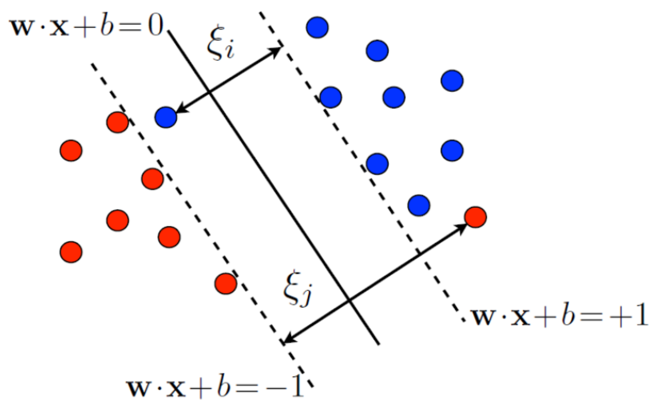
: 두 범주에 대해 선형분류를 최대한 잘 하는 초평면 (hyperplane)을 찾는 기법



- 직선을 그으면 두 범주를 완벽하게 분류하기 어렵기 때문에 SVM을 변형한 두가지 기법이 존재함

1) C-SVM

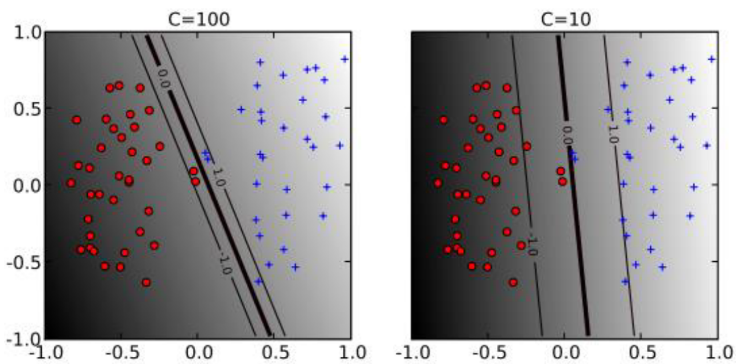
: minus-plane과 plus-plane사이에 관측지가 존재할 수 있도록 제약을 완화하는 방법



- 각 plane을 벗어난 빨간점과 파란점 같은 관측지들을 허용하되, 벗어난 ξ만큼 패널티 부과

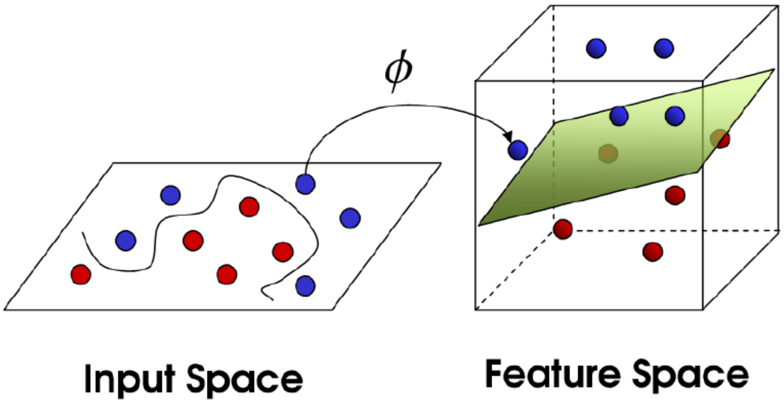
$$\min \frac{1}{2}\|w\|_2^2 + C\sum_{i=1}^{n}\xi_i$$

- 사용자가 설정하는 하이퍼파라미터인 C는 커질수록 허용하는 마진 폭이 줄어든다
- 원 SVM의 제약식: $y_i(w^T x_i + b) \geq 1$
- C-SVM의 제약식: $y_i(w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0$제
- 라그랑지안 접근법을 이용하여 해 도출
- C가 클수록 ξ_i의 역할이 커지고 그만큼 마진 폭이 줄어든다



2) Kernel-SVM

: 분류 경계면을 비선형 모양으로 만드는 SVM 기법. 원공간의 입력 데이터를 선형분류가 가능한 고차원 공간으로 매핑한 뒤 범주를 분류하는 초평면을 찾는게 목적



- **Mapping Function**: input space와 feature space 사이를 매핑해주는 함수
- **Kernel Trick**: 고차원 매핑과 내적을 한번에 하기 위해 도입된 것.

https://ratsgo.github.io/convex optimization/2018/01/25/duality/

https://datascienceschool.net/02 mathematics/05.02 제한조건이 있는 최적화 문제.html

https://ratsgo.github.io/machine learning/2017/05/23/SVM/