# SVM-KKT

SVM을 최적화 하기 위해선 KKT 조건을 활용한다.

SVM의 제약조건은 릿지회귀에서 보다 더 복잡하므로 KKT 조건을 사용하여 최적화를 진행한다.

목적함수 $\min \frac{1}{2}\|w\|^2$

제약조건 $y_i(w^T x_i + b) \ge 1$

라그랑주 함수 정의

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i \{ y_i (w^T x_i + b) - 1 \}$$

KKT 조건

① stationarity

$$\nabla L(w, b, \alpha) = 0$$

② primal feasibility

$$y_i (w^T x_i + b) \ge 1$$

③ dual feasibility

$$\alpha_i \ge 0$$

④ complementary slackness

$$\alpha_i \{ y_i (w^T x_i + b) - 1 \} = 0$$

Stationarity

$\nabla L(w, b, \alpha)$를 계산해보겠다.

$$L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i \{ y_i (w^T x_i + b) - 1 \}$$

$$\frac{d}{dw} L(w, b, \alpha) = w - \sum_{i=1}^{n} \alpha_i y_i x_i = 0$$

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i$$

$$\frac{d}{db} L(w, b, \alpha) = -\sum_{i=1}^{n} \alpha_i y_i = 0$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

종합하자면, 다음 2가지 결과를 얻었다

$$w^* = \sum_{i=1}^{n} \alpha_i y_i x_i$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

$L(w, b, \alpha)$를 $w^* = \sum_{i=1}^{n} \alpha_i y_i x_i$ 를 이용하여

dual 문제 $L(w^*, b^*, \alpha)$로 바꾸어 풀어보겠다.

$$\frac{1}{2}\|w^*\|^2 = \frac{1}{2} w^{*T} w^*$$

$$= \frac{1}{2} w^{*T} \sum_{i=1}^{n} \alpha_i y_i x_i$$

$$= \frac{1}{2} \sum_{i=1}^{n} \alpha_i y_i (w^{*T} x_i)$$

$$w^{*T} = \sum_{j=1}^{n} \alpha_j y_j x_j^T$$

$$= \frac{1}{2} \sum_{i=1}^{n} \alpha_i y_i \sum_{j=1}^{n} \alpha_j y_j x_j^T x_i$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_j^T x_i$$

$$\sum_{i=1}^{n} \alpha_i \left\{ y_i ( w^{*T} x_i + b ) - 1 \right\}$$

$$= \sum_{i=1}^{n} \alpha_i y_i ( w^{*T} x_i + b ) - \sum_{i=1}^{n} \alpha_i$$

$$= \sum_{i=1}^{n} \alpha_i y_i W^{*T} x_i + b \sum_{i=1}^{n} \alpha_i y_i - \sum_{i=1}^{n} \alpha_i$$

$$\sum_{i=1}^{n} \alpha_i y_i = 0$$

$$\sum_{i=1}^{n} \alpha_i y_i W^{*T} x_i = \|W^*\|^2$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_j^T x_i - \sum_{i=1}^{n} \alpha_i$$

이를 통해서 $L(w^*, b^*, \alpha) = \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_j^T x_i$

원래 라그랑주 함수는 $L(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i \{ y_i (w^T x_i + b) - 1 \}$

목적함수인 $\frac{1}{2}\|w\|^2$을 최소화하려면 $\alpha$를 최대화 해야 한다.

$L(w^*, b^*, \alpha)$를 $L_{dual}(\alpha)$로 둔다면 이를 최대화 해야 한다.

정리하자면,

$$\text{maximize } L_{dual}(\alpha) = \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_j^T x_i$$

$$\text{subject to } \begin{cases} \sum_{i=1}^{n} \alpha_i y_i = 0 \\ \alpha_i \geq 0 \end{cases}$$