# Двойственность

## Условие ККТ

$$\min f(x) = p^*$$
$$s.t.\ g_i(x) \le 0$$
$$h_i = 0$$

$$L(x, \lambda, \nu) = f(x) + \lambda^T h(x) + \nu^T g(x)$$

1) $\nabla_x L = 0$

2) $\nabla_\lambda L = 0$

3) $\nu^* \ge 0$

4) $g(x^*) \le 0$

5) $\nu^* g(x^*) = 0$

Лагранжиан по построению всегда меньше

$$g(\lambda, \nu) = \inf_x L(\nu, \lambda, x) = \inf_x f(x) + \lambda^T h(x) + \nu^T g(x)$$

$$g(\lambda, \nu) \le p^*$$

$$\lambda^T h(x) + \nu^T g(x) \le 0$$

$$L(x, \lambda, \nu) = f(x) + \lambda^T h(x) + \nu^T g(x) \le 0$$

$$g(\lambda, \nu) \le L(x, \nu, \lambda) \le p^*$$

# Дуальная функция Лагранжиана

$$g(\lambda, \nu) = \inf L$$

- Вогнутая
- $g \le p^*$

## Двойственная задача

$$\max_{\lambda, \nu} g(\lambda, \nu)$$
$$\text{s.t. } \nu \ge 0$$

**Пример.** Построить двойственную

$$\min \|x\|_2^2$$
$$\text{s.t. } Ax = b$$

$$L(x, \lambda) = x^T x + \lambda^T (Ax - b)$$

$$\nabla_x L = 2x + A^T \lambda \qquad x = -\frac{1}{2} A^T \lambda$$

$$g(\lambda) = L\left(\lambda, -\frac{1}{2} A^T \lambda\right) = -\frac{1}{4} \lambda^T A A^T \lambda - b^T \lambda$$

$$\max -\frac{1}{4} \lambda^T A A^T \lambda - b^T \lambda$$

Пример. Построить двойственную задачу

$$\begin{aligned} \min\ & c^T x \\ \text{s.t. } & Ax = b \\ & x \geq 0 \end{aligned}$$

$\nu \quad \lambda$

$$L(x, \nu, \lambda) = c^T x - \nu^T x + \lambda^T (Ax - b) = 0$$

$$L(x, \nu, \lambda) = -\lambda^T b + (c - \nu + A^T \lambda)^T x = 0$$

$\inf L$ : ?

$$\inf L = -\lambda^T b + \inf_x (c - \nu + A^T \lambda)^T x$$

$$g(\lambda, \nu) = \begin{cases} -\lambda^T b & A^T\lambda - \nu + c = 0 \quad \nu \geq 0 \\ -\infty & \text{otherwise} \end{cases}$$

$$\begin{aligned} \max\ & -\lambda^T b \\ \text{s.t. } & A^T\lambda - \nu + c = 0 \\ & \nu \geq 0 \end{aligned}$$

Пример: Невыпуклая задача.

$$\min x^T W x$$
$$\text{s.t. } x_i^2 = 1$$

$$L(x, \lambda) = x^T W x + \lambda^T (x^T x - 1) =$$
$$= x^T (W + \mathrm{diag}\,\lambda) x - 1^T \lambda$$

$$g(\lambda) = \inf_x \left( x^T (W + \mathrm{diag}\,\lambda) x \right) - 1^T \lambda$$

$$g(\lambda) = \begin{cases} -1^T \lambda & x^T (W + \mathrm{diag}\,\lambda) x \succeq 0 \\ -\infty & \end{cases}$$

# Двойственная задача

$$\max g(\lambda, \nu) = d^*$$
$$\text{s.t. } \nu \succeq 0$$

Как соотносятся решения прямой и двойст задач?

$$d^* \le p^*$$

Зазор двойственности $d^* - p^*$

$p^* = d^*$ сильная двойственность

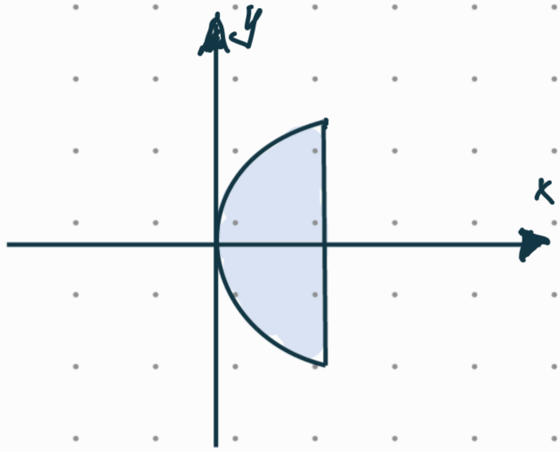
$p^* > d^*$ слабая двойственность есть всегда

## Условие Слейтера (достаточное)

Если $x^*$ лежит строго внутри допустимого мн-ва, то есть сильная двойственность

Задача выпукла и есть Слейтер $\to$ KKT

Задача выпукла + KKT $\to$ сильная двойственность




$$A = \{ (x,y) \; 0 \le x \le 1, \; |y| \le \sqrt{x} \}$$

$$\min y$$
$$\text{s.t } x \le 0$$

- The student is following some new diet trend which requires her to eat at least 6oz of chocolate, 8oz of cream cheese, and 10oz of sugar.
- Her goal is to satisfy these requirements at minimal cost.

| Ingredients needed |  |  |  |  |
|---|---|---|---|---|
|  | 3 oz | 2 oz | 2 oz | 50 cts |
|  | 0 oz | 4 oz | 5 oz | 80 cts |
| Requirements | 6 oz | 10 oz | 8 oz | |

$$\min_x \ 50x_1 + 80x_2$$

$$3x_1 \geq 6$$
$$2x_1 + 4x_2 \geq 10$$
$$2x_1 + 5x_2 \geq 8$$
$$x_1, x_2 \geq 0$$

$$\max \ 6y_1 + 10y_2 + 8y_3$$

$$3y_1 + 2y_2 + 2y_3 \leq 50$$
$$4y_2 + 5y_3 \leq 80$$
$$y_1, y_2, y_3 \geq 0$$

$$L = 50x_1 + 80x_2 + \nu_1(6 - 3x_1) + \nu_2(10 - 2x_1 - 4x_2) + \nu_3(8 - 2x_1 - 5x_2)$$

$$L = (50 - 3\nu_1 - 2\nu_2 - 2\nu_3)x_1 + (80 - 4\nu_2 - 5\nu_3)x_2 + 6\nu_1 + 10\nu_2 + 8\nu_3$$

$$\inf L = 6\nu_1 + 10\nu_2 + 8\nu_3 + \inf_{x_1 \geq 0}[\overset{a}{(50 - 3\nu_1 - 2\nu_2 - 2\nu_3)}x_1] + \inf_{x_2 \geq 0}[\overset{b}{(80 - 4\nu_2 - 5\nu_3)}x_2]$$

$$\inf a \begin{cases} 0 & a \geq 0 \\ -\infty \end{cases}$$

$$\inf b \begin{cases} 0 & b \geq 0 \\ -\infty \end{cases}$$

$$\max \ 6\nu_1 + 10\nu_2 + 8\nu_3$$
$$50 \geq 3\nu_1 + 2\nu_2 + 2\nu_3$$
$$80 \geq 4\nu_2 + 5\nu_3$$
$$\nu_1 \quad \nu_2 \quad \nu_3$$

дв задача

# Метод опорных векторов

У нас есть линейный классификатор

$$a(x,w) = \operatorname{sign}(\langle w, x\rangle - w_0)$$

и линейно разделимая выборка $X^l = (x_i, y_i)^l$

$\Rightarrow w, w_0 \quad M_i(w, w_0) = y_i(\langle w, x_i\rangle - w_0) \geq 0$

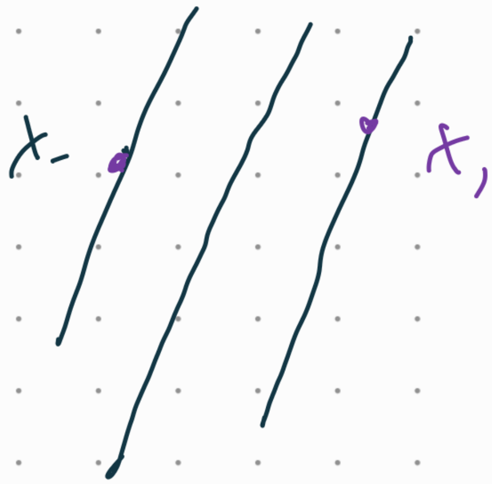
нормировка $\min M_i(w, w_0) = 1$

Разделяющая полоса (гиперплоскость)

$\{x : -1 \le \langle w, x_0\rangle - w_0 \le 1\}$

$\exists x_+ : \langle w, x_+\rangle - w_0 = +1$

$\exists x_- : \langle w, x_-\rangle - w_0 = -1$




Надо провести гиперплоскость так чтобы разд. полоса была шире

Ширина полосы

$$\frac{\langle x_+ - x_-, w\rangle}{\|w\|} = \frac{2}{\|w\|} \to \max$$

$$\Rightarrow \begin{cases} \frac{1}{2}\|w\|^2 \to \min \\ M_i(w, w_0) \geq 1 \end{cases}$$

$$\begin{cases} \frac{1}{2}\|w\|^2 \to \min \\ M_i(w, w_0) \geq 1 \end{cases}$$ линейно разделимая выборка.

В общем случае система несовместна переходим к линейно не разделимой выборке

$$\begin{cases} \frac{1}{2}\|w\|^2 + C\sum \xi_i \to \min \\ M_i(w, w_0) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases}$$

$$\Downarrow$$

$$\begin{cases} \frac{1}{2}\|w\|^2 + C\sum \xi_i \to \min \\ -M_i(w, w_0) + 1 - \xi_i \leq 0 \quad \longleftrightarrow \lambda_i \\ -\xi_i \leq 0 \quad \longleftrightarrow \quad \eta_i \end{cases}$$

$$L(w, w_0, \xi, \lambda) = \frac{1}{2}\|w\|^2 + C\sum \xi_i - \sum \eta_i \xi_i - \sum (M_i - 1 + \xi_i)\lambda_i =$$

$$= \frac{1}{2}\|w\|^2 - \sum \lambda_i (M_i - 1) - \sum \xi_i (\lambda_i + \eta_i - C)$$

Прямые переменные $w, w_0, \xi$

$$\frac{\partial L}{\partial w} = 0 \qquad \frac{\partial L}{\partial w_0} = 0 \qquad \frac{\partial L}{\partial \xi} = 0$$

$$-\xi_i \le 0 \qquad \lambda_i \geq 0 \qquad \eta_i \geq 0$$

$$\lambda_i = 0 \quad \text{or} \quad M_i = 1 - \xi_i$$

$$\eta_i = 0 \quad \text{or} \quad \xi_i = 0$$

$$\frac{\partial L}{\partial w} = w - \sum \lambda_i y_i x_i = 0 \qquad w = \sum \lambda_i y_i x_i$$

0 $$\frac{\partial L}{\partial w_0} = -\sum y_i \lambda_i = 0 \qquad \sum y_i \lambda_i = 0$$

$$\frac{\partial L}{\partial \xi_i} = -\lambda_i - \eta_i + C = 0 \qquad C = \lambda_i + \eta_i$$

Объекты можно разделать на три типа

1) $\lambda_i = 0 \quad \eta_i = C \Rightarrow \xi_i = 0 \quad M_i \geq 1$

периферийные объекты

2) $\lambda_i \in (0, C), \ \eta_i \in (0, C) \quad \xi_i = 0 \ M_i = 1$

опорные граничные объекта

3) $\lambda_i = C \quad \eta_i = 0 \quad \xi_i \geq 0 \quad M_i < 1$

Опорные нарушители

Объект опорный $\lambda_i \neq 0$

$$L = \frac{1}{2}\|w\|^2 - \sum \lambda_i (M_i - 1) - \sum \xi_i (\nu_i + \eta_i - C)$$

Распишем лагранжиан для опорных объектов

$$\begin{cases} L = -\frac{1}{2}\sum \lambda_i \lambda_j y_i y_j x_i x_j + \sum_{i=1}^{l} \lambda_i \\ 0 \le \lambda_i \le C \\ \sum \lambda_i y_i = 0 \end{cases}$$

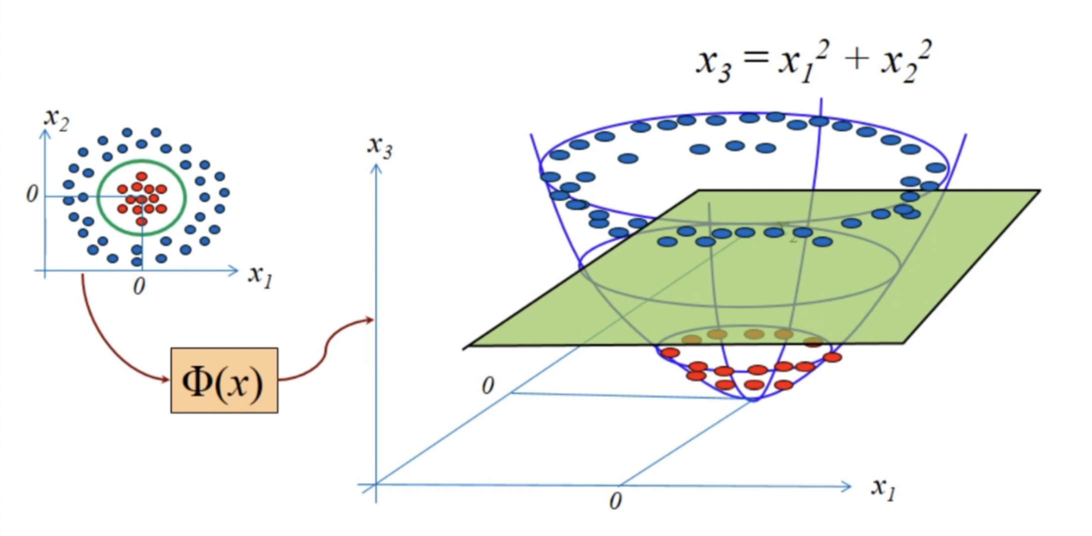
выпуклая задача

Решение прямой выражается через решение двойственной

$$\begin{cases} w = \sum \lambda_i y_i x_i \\ w_0 = \langle w, x_i \rangle - y_i \end{cases} \quad \lambda_i > 0 \quad M_i = 1$$

$$a(x) = \mathrm{sign}\left(\sum \lambda_i y_i \langle x, x_i \rangle - w_0\right)$$

Скалярное произведение можно заменить на ядро $k(\langle x, x \rangle)$




$$k(u,v) = \langle u, v\rangle^2 \qquad u(u_1, u_2) \quad v(v_1, v_2)$$

$$k(u,v) = \langle u,v\rangle\langle u,v\rangle = (u_1 v_1 + u_2 v_2)^2 =$$

$$= (u_1 v_1)^2 + (u_2 v_2)^2 + 2u_1 v_1 u_2 v_2 =$$

$$= \langle (u_1^2, u_2^2, \sqrt{2}u_1 u_2)(v_1^2, v_2^2, v_1 v_2 \sqrt{2})\rangle$$