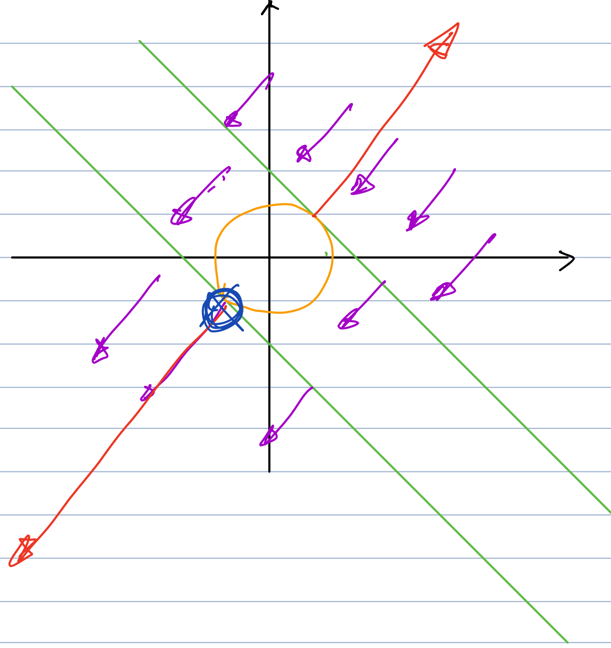
拉格朗日乘子法:

① $\min f(x)$

$\nabla f(x) = 0$

② $\begin{cases} \min f(x) \\ s.t \quad h(x) = 0 \end{cases}$

例: $\begin{cases} \min x_1 + x_2 \\ x_1^2 + x_2^2 = 2 \end{cases}$



$f(x) = x_1 + x_2$, $\quad h(x) = x_1^2 + x_2^2 - 2$

$-\nabla f(x) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$, $\quad \nabla h(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$

由图示可知 $-\nabla f(x)$, $\nabla h(x)$ 夹角越小，越逼近最小值

即最小值为 $\nabla f(x) = -\lambda \nabla h(x)$, $h(x) = 0$

整合一下有 $\begin{cases} \min f(x) + \lambda h(x) \\ \lambda \geq 0 \end{cases}$

③ $\begin{cases} \min f(x) \\ g(x) \leq 0 \end{cases}$ $\Rightarrow$ (1) 最优点在 $g(x) < 0$ 内，相当没有约束 $\lambda = 0$

(2) 最优点在 $g(x) = 0$ 即边界上。

$\Rightarrow \begin{cases} \min f(x) + \lambda g(x) \\ \lambda \geq 0 \\ \lambda g(x) = 0 \\ g(x) \leq 0 \end{cases}$

$\begin{cases} \min x^2 \\ x \geq b \end{cases}$

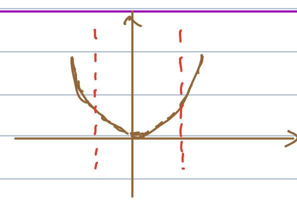
$\Downarrow$

$\begin{cases} \min x^2 \\ b - x \leq 0 \end{cases}$

$\begin{cases} L = x^2 + \lambda(b-x) \\ \lambda \geq 0 \end{cases}$

$\Downarrow$

$\begin{cases} \frac{\partial L}{\partial x} = 0 \Rightarrow 2x - \lambda = 0 \Rightarrow x = \frac{\lambda}{2} \\ \frac{\partial L}{\partial \lambda} = 0 \Rightarrow x = b \\ \lambda \geq 0 \\ \lambda(b-x) = 0 \Rightarrow \lambda = 0 \text{ 或 } x = b \end{cases}$



$\min L = \max(0, b)$

举个例子，可以忽略。

总结

$\begin{cases} \min f(x) \\ s.t \quad h_i(x) = 0 \quad i = 1, \cdots, l \\ \quad\quad g_j(x) \leq 0, \quad j = 1, \cdots, m \end{cases}$ $\Rightarrow$ $L = \min f(x) + \sum_{i=1}^{l} \alpha_i h_i(x) + \sum_{j=1}^{m} u_j g_j(x)$

$\begin{cases} \frac{\partial L}{\partial x} = 0, \quad \frac{\partial L}{\partial \alpha_i} = 0, \quad \frac{\partial L}{\partial u_j} = 0 \\ \alpha_i \geq 0, \quad u_j \geq 0 \\ u_j g_j(x) = 0 \\ \cdot(x) \leq 0 \end{cases}$

关于对偶问题：

原式
$$\begin{cases} \min f(x) \\ s.t \quad h_i(x)=0, \quad i=1,\dots,l \\ \quad g_j(x)\le 0, \quad j=1,\dots,m \end{cases}$$

最优值 p^*

$$L = \min f(x) + \sum_{i=1}^{l} \alpha_i h_i(x) + \sum_{j=1}^{m} \nu_j g_j(x)$$

最优值 d^*

$\because$ L 中 $\nu_j, \alpha_j \geq 0$, $g_j(x) \le 0$ $\therefore$ $d^* \le p^*$ 此时为弱对偶问题.

但有 $f(x), g(x)$ 为凸函数 $h(x)$ 为仿射函数 有 $d^* = p^*$ 为强对偶问题.

此时 max, min 位置可以互换.

SMO：1. 启发式找 a_i, a_j

2. 固定 $a_i \cdot a_j$ 以外变量，求 a_i, a_j

$a_i y_i + a_j y_j = C$, $a_i \geq 0$, $a_j \geq 0$

$$a_j = \frac{C - a_i y_i}{y_i}$$ ，变成单变量优化

$$b = \frac{1}{|S|}\left(\frac{1}{y} - \sum_{x \in S} a_i y_i x_i^T x_s \right)$$

问题

6.4 LDA, SVM 两者 $w_1 * w_2 = 0$ 时等价

6.5 不清楚

6.6 离群越远，拉格朗日乘子部分越大

6.9 $w = \sum_{i=1}^{m} a_i \phi(x_i)$ $w = \sum_i a_i * \varphi(x_i)$

$$z = w^T x + b = \sum_{i=1}^{m}\sum_{j=1}^{m} a_i \phi(x_j)^T \phi(x_0) + b$$
$$= \sum_{i=1}^{m} a_i k(x, x_i) + b$$

$$y = \frac{1}{1+e^{-z}}$$

6.10 取 α 的 tip 5