lesson 36~38 KKT条件

KKT条件：

原问题： $\min f_0(x)$ p^*

s.t. $f_i(x) \le 0$ $i=1,2\cdots m$

$h_i(x)=0$ $i=1,2\cdots p$

对偶问题： $\sup\limits_{\lambda\ge0} \inf\limits_{x\in D} L(x,\lambda,v)$

$$= \sup_{\lambda\ge0} \inf_{x\in D} \left(f_0(x)+\sum_{i=1}^{m}\lambda_i f_i(x)+\sum_{i=1}^{p} v_i h_i(x)\right) \quad d^*$$

引入假设： 1. $d^*=p^*$，对偶间隙为0

2. 所有的 f, h 都是可微的

推导： 假设最优解为：(x^*, λ^*, v^*)

则
$$\begin{cases} f_i(x^*) \le 0 & i=1,2\cdots m \\ h_i(x^*)=0 & i=1,2\cdots p \\ \lambda^* \ge 0 \end{cases}$$

原问题的可行域（前两行）

对偶问题的可行域（第三行）

$$p^* = d^*$$

$$\Rightarrow f_0(x^*) = \inf_x L(x,\lambda^*,v^*)$$

$$= \inf_x \left(f_0(x)+\sum_{i=1}^{m}\lambda_i^* f_i(x)+\sum_{i=1}^{p} v_i^* h_i(x)\right)$$

1)

$$\le f_0(x^*)+\sum_{i=1}^{m}\lambda_i^* f_i(x^*)+\sum_{i=1}^{p} v_i^* h_i(x)$$

2)

$$\le f_0(x^*)$$

由(2)可知:

$$\sum_{i=1}^{m} \underbrace{\lambda_i^*}_{\geq 0} \underbrace{f_i(x^*)}_{\leq 0} = 0$$

$\lambda_i^* f_i(x^*) = 0, \ \forall i=1,2\cdots m$

$\begin{cases} f_i(x^*) < 0, & 则\ \lambda_i = 0 \\ \lambda_i > 0, & 则\ f_i(x^*) = 0 \end{cases}$

互补松弛条件

由(1)可知:

$$x^* = \operatorname{argmin} L(x, \lambda^*, \nu^*)$$

则: $$\left.\frac{\partial L(x, \lambda^*, \nu^*)}{\partial x}\right|_{x=x^*} = 0$$

$$\Rightarrow \nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0$$

稳定性条件!

KKT条件:

① 原问题的可行性

② 对偶问题的可行性

③ 互补松弛条件

④ 稳定性条件

KKT条件的充要性：

若原问题是凸问题，各个函数可微且对偶间隙为0，

则KKT条件为充要条件

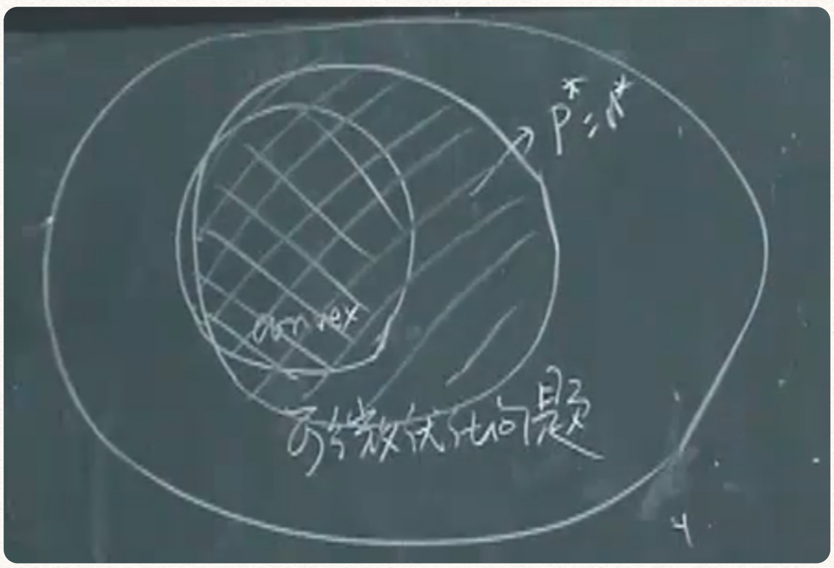
证明充分性：

即证明满足KKT条件的$(\tilde{x}, \tilde{\lambda}, \tilde{v})$是原问题的最优解.

即证： $g(\tilde{\lambda}, \tilde{v}) = f(\tilde{x})$

$$
\begin{aligned}
g(\tilde{\lambda}, \tilde{v}) &= \inf_{x} L(x, \tilde{\lambda}, \tilde{v}) \\
&= L(\tilde{x}, \tilde{\lambda}, \tilde{v}) \\
&= f_0(\tilde{x}) + \sum_{i=1}^{m} \underbrace{\tilde{\lambda}_i f_i(\tilde{x})}_{=0} + \sum_{i=1}^{p} \underbrace{\tilde{v}_i h_i(\tilde{x})}_{=0} \\
&= f_0(\tilde{x})
\end{aligned}
$$

相关问题之间的关系：



一般的优化问题：（不一定凸）

若拉格朗日函数有鞍点 $\Leftrightarrow$

此时 p/d 最优对偶间隙为0

可微的优化问题：

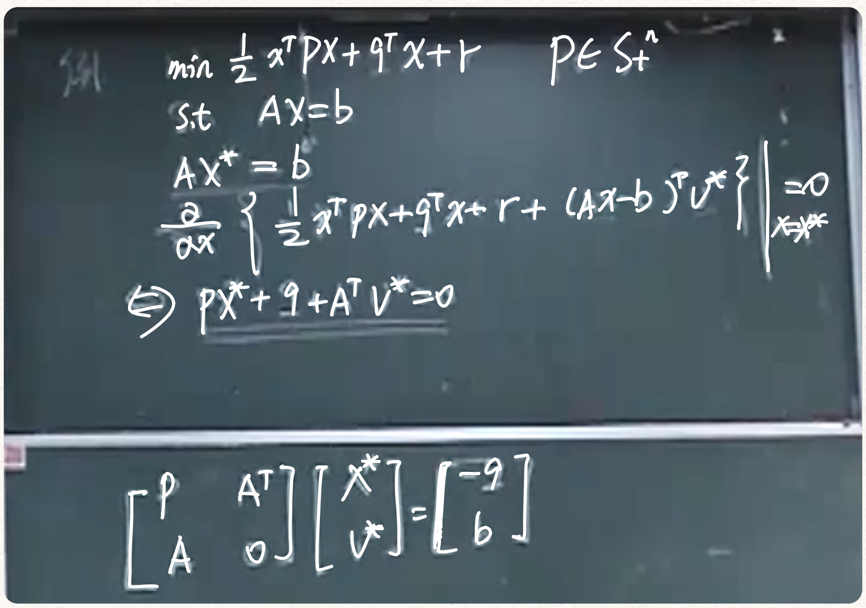
对偶间隙为0的必要条件为KKT条件：

$$\begin{aligned}\min\ & f_0(x)\\ \text{s.t.}\ & f_i(x)\le 0 \quad i=1\cdots m\\ & h_i(x)=0 \quad i=1\cdots p\end{aligned}\quad \xRightarrow{\text{KKT}} \quad \begin{cases} f_i(x^*)\le 0\\ h_i(x^*)=0\\ \lambda^*\ge 0\\ \lambda_i^* f_i(x^*)=0\\ \nabla f_0(x^*)+\sum_{i=1}^{m}\lambda_i^*\nabla f_i(x^*)+\sum_{i=1}^{p}\nu_i^*\nabla h_i(x^*)=0\end{cases}$$

一般可微的凸优化问题：

KKT条件为对偶间隙为0的充要条件！

例：二次函数KKT条件



例: $x\geq 0$ 约束下的最优化

$$\min \quad f_0(x)$$
$$\text{s.t.} \quad x \geq 0$$

互补条件!

$$\begin{cases} x \geq 0 \\ Df_0(x) \geq 0 \\ x_i \left(Df_0(x)\right)_i = 0 \end{cases}$$

KKT条件:

$$\begin{cases} x^* \geq 0 & ① \\ \lambda^* \geq 0 & ② \\ \lambda_i^* \cdot (-x_i^*) = 0 & ③ \\ Df_0(x^*) - \lambda^* = 0 & ④ \end{cases}$$

由②和④得:

$$Df_0(x^*) = \lambda^* \geq 0$$

由③和④得:

$$\left(Df_0(x^*)\right)_i \cdot x_i = 0$$