优化理论——拉格朗日乘子法和KKT条件

星空爱好者
优化理论.机器学习

导语

拉格朗日乘子法和KKT条件应用于求解带约束条件的优化模型
拉格朗日乘子法的意义：通过引入新的自由变量，将等式不等式的约束条件转化为无约束或约束简单的优化问题进行求解。

KKT条件的意义：满足KKT条件的K-T点是模型最优解的必要条件，对于凸优化而言，K-T点是最优解的充要条件。通常直接求解优化模型往往比较困难，转而求解KKT条件方程组的解，再去验证方程解是否为模型的最优解，从而大幅度提升了优化模型的求解效率。

等式约束模型

对于等式约束**原模型**：

$$
\begin{aligned}
&min\ f(x)\\
&s.t.\ g(x)=c\\
&\qquad x\in R
\end{aligned}
$$

引入自由变量 λ，构建拉格朗日函数 $L(x,\lambda)=f(x)+\lambda(g(x)-c)$，接下来证明原模型等价于无约束优化模型

$$
\begin{aligned}
&min\ L(x,\lambda)\\
&s.t.\ x,\lambda\in R
\end{aligned}
$$

等价证明1

拉格朗日函数对偶形式为

$\max_\lambda \theta_d(\lambda)=\max_\lambda \min_x f(x)+\lambda(g(x)-c)$

假设 $x^*=\arg\min_x f(x)+\lambda(g(x)-c)$

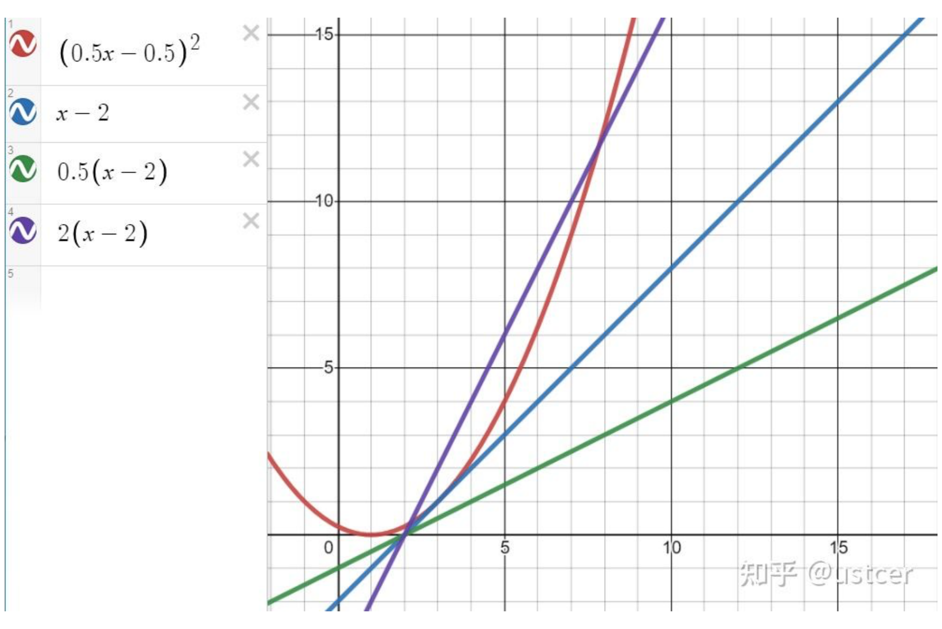
则 $\max_\lambda\ L(x^*,\lambda)$，根据约束分为两种情况
$$
\begin{cases} f(x^*), & g(x^*)=c\\ f(x^*)+\lambda(g(x^*)-c), & g(x^*)\neq c\end{cases}
$$

λ是自由变量，$g(x^*)\neq c$ 时为线性函数，可以取到无穷大或者无穷小，$\max_\lambda\ L(x^*,\lambda)$ 只能取 $g(x^*)=c$ 的第一种情况。因此$min\ L(x,\lambda)$ 等价于原模型。

等价证明2
模型存在最优解，则目标函数 $f(x)$ 与约束函数 $g(x)=c$ 必然相交。凸函数最优点只有一个，所以只有约束条件与目标函数等高线的切点才是优化模型的最优解。切点处，目标函数梯度与约束函数梯度平行，公式化为

$$f'(x)=-\lambda g'(x)$$

从而验证了拉格朗日函数 $L(x,\lambda)=f(x)+\lambda(g(x)-c)$的正确性。



\labmda = 1时, h(x)和g(x)相切，切点为最优点；\labmda = 2时, h(x)和g(x)相交；\labmda = 0.5时, h(x)和g(x)相离无交点。寻找最优点变成了调整\lambda值寻找切点

等式约束下的KKT条件

这里的拉格朗日函数 $L(x,\lambda)=f(x)+\lambda(g(x)-c)$已经成为了无约束函数，最优解自然满足其在各个自由变量处的导数为0。因此，等式约束优化模型的KKT条件为
$$
\begin{cases} \frac{\partial L}{\partial x}=f'(x)+\lambda g'(x)=0\\ \frac{\partial L}{\partial \lambda}=g(x)-c=0\end{cases}
$$

不等式约束模型

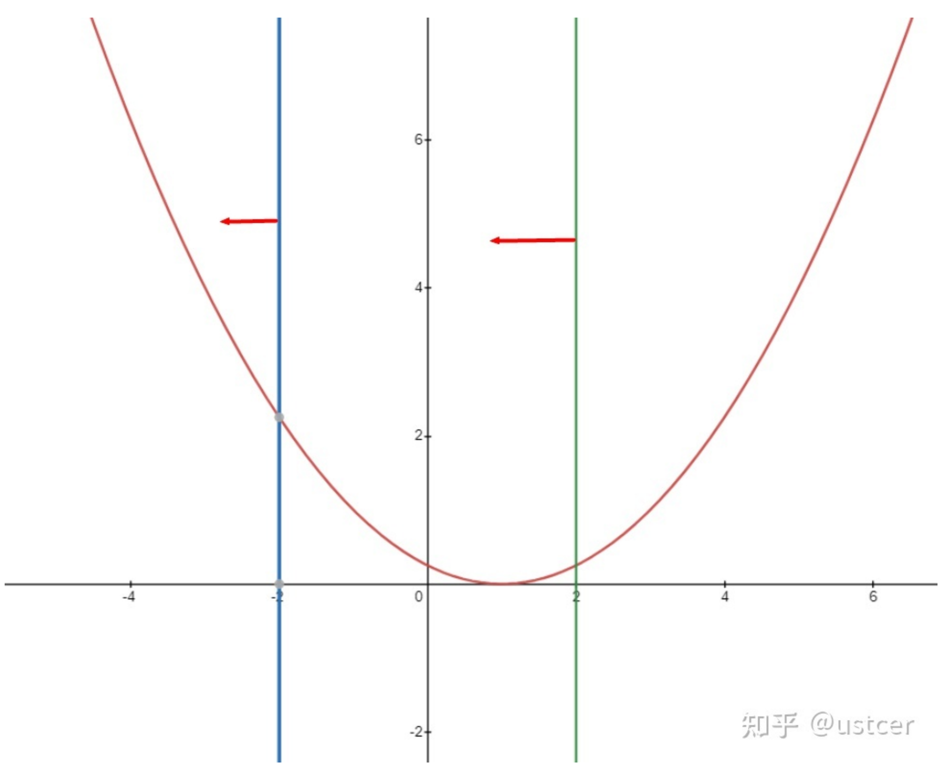
对于不等式约束**原模型**：
$$
\begin{aligned}
&min\ f(x)\\
&s.t.\ h(x)\le t\\
&\qquad x\in R
\end{aligned}
$$

这里引入一个自由变量 s 使得不等式约束问题转化成等式约束问题如下
$$
\begin{aligned}
&min\ f(x)\\
&s.t.\ h(x)+s^2=t\\
&\qquad x,s\in R
\end{aligned}
$$

对应的拉格朗日函数 $L(x,\mu,s)=f(x)+\mu(h(x)+s^2-t)$，等式约束优化模型的KKT条件为

$$
\begin{cases} \frac{\partial L}{\partial x}=f'(x)+\mu h'(x)=0\\ \frac{\partial L}{\partial \mu}=h(x)+s^2-t=0\\ \frac{\partial L}{\partial s}=2\mu s=0\end{cases}
$$

首先由于KKT条件要求 $f'(x)=-\mu h'(x)$ 。对于目标函数，约束起作用意味着越是远离约束边界的解空间使得目标函数值越大，t 已经属于$h(x)$ 最大取值，随着可行解远离约束边界 $h(x)$ 递减。这两个函数的梯度是相反的，因此 $f'(x)$ 与$h'(x)$ 异号，所以拉格朗日乘子 $\mu\ge 0$ 。此处解释可能较为抽象，以下图为例更容易理解。



上图中，蓝线约束 $x\le -2$ 为有效约束，最优解在约束边界取得(左交点)，目标函数（红线）随着 x 远离边界而递增，可行域内约束条件的梯度大于0而目标函数梯度小于0。绿色约束 $x\le 2$ 为无效约束，因为最优解在为$x=1$。

接下来重点分析下 $\mu s=0$，分为两种情况。

第一种情况：$\mu=0,s\neq 0$，$h(x^*)<t$
最优解 x^* 在不等式内部，此时不等式约束无效，即拿掉此不等式约束对最优解无影响

第二种情况：$\mu\ge 0,s=0$，$h(x^*)=t$
最优解在不等式约束的边界上，此时不等式约束转化为等式约束。
基于对 $\mu s=0$ 的讨论，等价于 $\mu(h(x)-t)=0$

综上，不等式模型的KKT条件总结如下：

$$
\begin{cases} f'(x)+\mu h'(x)=0\\ \mu(h(x)-t)=0\\ h(x)-t\le 0\\ \mu\ge 0\end{cases}
$$

等式不等式约束模型

对于等式不等式联合约束的优化模型

$$
\begin{aligned}
&min\ f(x)\\
&s.t.\ g(x)=c\\
&\qquad h(x)\le t\\
&\qquad x\in R
\end{aligned}
$$

此处不再赘述，给出拉格朗日函数 $L(x,\lambda,\mu)=f(x)+\lambda(g(x)-c)+\mu(h(x)-t)$，KKT条件为

$$
\begin{cases} f'(x)+\lambda g'(x)+\mu h'(x)=0\\ g(x)=c\\ h(x)\le t\\ \mu(h(x)-t)=0\\ \mu\ge 0\end{cases}
$$

编辑于 2022-05-07 18:40

凸优化 不等式 拉格朗日乘子