#### 1.Lagrange对偶函数

* **Lagrange**

考虑标准形式的优化问题



其中自变量，设问题的定义域是非空集合，优化问题的最优值为，之一这里并没有假设是凸优化问题。

Lagrange对偶的基本思想是在目标函数中考虑问题的约束条件，即添加约束条件的加权和，得到增广的目标函数。定义问题的Lagrange函数为



其中定义域为。称为第个不等式约束对应的**Lagrange乘子**；类似地，称为第个等式约束对应的Lagrange乘子。向量和称为对偶变量或者是问题的**Lagrange乘子向量**。

* **Lagrange对偶函数**

定义**Lagrange对偶函数**（或简称**对偶函数**）为Lagrange函数关于取得的最小值：即对，，有



如果Lagrange函数关于无下界，则对偶函数取值为。因为对偶函数是一族关于的仿射函数的逐点下确界，所以即使原问题不是凸的，对偶函数也是凹函数。

* **最优值的下界**

对偶函数构成了原问题最优值的下界：即对任意和下式成立



这是由于每个可行点都满足，因此不等式成立。但当时其意义不大。只有当且，即时，对偶函数才能给出的一个非平凡下界。称满足以及的是**对偶可行**的。

*该部分可以参考<https://www.zhihu.com/question/58584814>的解释*

* Lagrange对偶函数和共轭函数

第三章中提到的函数的共轭函数为



事实上，共轭函数和Lagrange对偶函数密切相关。考虑优化问题



利用函数的共轭函数，可以将以上问题的对偶函数表述为



函数的定义域也可以由函数的定义域得到



#### 2.Lagrange对偶问题

Lagrange对偶函数给出了优化问题的最优值的一个下界，那么其能够得到的最好下界是什么？这个问题可以表述为优化问题



上述问题称为**Lagrange对偶问题**。在满足对偶可行（即和）时，对偶问题存在最优解，称为**对偶最优解**或者**最优Lagrange乘子**。

Lagrange对偶问题是一个凸优化问题，这是因为极大化的目标函数是凹函数，且约束集合是凸集。因此，对偶问题的凸性和原问题是否是凸优化问题无关。

* 显式表达对偶约束

很多情况下我们可以求出对偶函数定义域的仿射包并将其表示为一系列线性等式约束，这样处理后我们可以得到一个等价的问题。

* **弱对偶性**

Lagrange对偶问题的最优值，我们用表示。根据定义，这是通过Lagrange函数得到的原问题最优值的最好下界，因此有



即使原问题不是凸问题，上述不等式亦成立。这个性质称为**弱对偶性**。定义差值是原问题的**最优对偶间隙**。

* **强对偶性和Slater约束准则**

如果等式



成立，那么强对偶性成立。这说明从Lagrange对偶函数得到的最好下界是紧的。

对于一般情况，强对偶性不成立。如果原问题是凸问题，强对偶性通常（但不总是）成立。许多研究成果给出了强对偶性成立的条件（除了凸性条件以外）。这些条件称为**约束准则**。

一个简单的约束准则是**Slater条件**：存在一点使得下式成立



满足上述条件的点有时称为**严格可行**，这是因为不等式约束严格成立。Slater定理说明，当Slater条件成立（且原问题是凸问题）时，强对偶性成立。

当不等式约束函数中有一些是仿射函数时，Slater条件可以进一步改进。如果最前面的个约束是仿射的，则若下列弱化的条件成立，强对偶性成立。该条件为：存在一点使得



换言之，仿射不等式不需要严格成立。

#### 3.几何解释

* 通过函数值集合理解强弱对偶性

可以通过集合



来给出对偶函数的简单几何解释。事实上，此集合是约束函数和目标函数所取得的函数值。显然满足关系



如果且，则成立，因此有



#### 5.最优性条件

* **次优解认证和终止准则**

如果能够找到一个对偶可行解，就对原问题的最优值建立了一个下界：。对偶可行点为表达式的成立提供了一个**证明**或**认证**。

对偶可行点可以让我们在不知道的确切值的情况下界定给定可行点的次优程度。事实上，如果是原问题的可行解且对偶可行，那么



特别地，上式说明了是ε-次优，其中。

定义原问题和对偶问题目标函数的差值



为原问题的可行解与对偶可行解之间的**对偶间隙**。对偶间隙将原问题（和对偶问题）的最优值限制在了一个区间上：



如果对偶间隙为0，那么是原问题的最优解并且是对偶问题的最优解。此时我们可以认为是证明为最优解的一个认证。

上述现象可以用在优化算法中给出非启发式停止准则。设某个算法给出一系列原问题可行解以及对偶问题可行解，给定要求的绝对精度，那么停止准则（即算法终止的条件）



保证大概算法终止的时候，是-次优。

类似地，我们可以定义相对精度来保证相对误差。

* **互补松弛性**

设原问题和对偶问题的最优值都可以达到且相等（即强对偶性成立）。令是原问题的最优解，是对偶问题的最优解，这表明



不难发现，上面的式子链中的两个不等式实际上都取等号。由此可以得出一些有意义的结论，例如，由第三个不等式变为等式，我们知道关于求极小时在处取得最小值。

另外一个重要的结论是



事实上，求和项的每一项都非正，因此有



上述条件称为互补松弛性。我们可以将互补松弛条件写成



或者等价地



* KKT最优性条件

现在假设函数可微（因此定义域是开集）。

非凸问题的KKT条件

和前面一样，令和分别是原问题和对偶问题的最优解，对偶间隙为0。因为关于求极小在处取得最小值，因此函数在处的导数必须为0，即



因此，我们有



我们称上式为**Karush-Kuhn-Tucker**（**KKT**）**条件**。

总之，对于目标函数和约束函数可微的任意优化问题，如果强对偶性成立，那么任何一对原问题最优解和对偶问题最优解必须满足KKT条件。

凸问题的KKT条件

当原问题是凸问题时，满足KKT条件的点也是原、对偶最优解。

KKT条件在优化领域有着重要的作用。在一些特殊的情形下，是可以解析求解KKT条件的（也因此可以求解优化问题）。更一般地，很多求解凸优化问题的方法可以认为或者理解为求解KKT条件的方法。

#### 6.扰动及灵敏度分析

当强对偶性成立时，对原问题的约束进行扰动，对偶问题最优变量为原问题最优值的灵敏度分析提供了很多有用的信息。

* 扰动的问题

考虑对原优化问题进行扰动之后的问题



当以及时，上述问题即为原问题。若，则我们放松了第个不等式约束；当，则意味着我们加强此约束。

定义为扰动的问题的最优值。有可能，这时对约束的扰动使得扰动后的问题不可行。

* 一个全局不等式

假设强对偶性成立且对偶问题的最优值可以达到，是未被扰动的问题的对偶问题最优解。对于所有的和，我们有



* 局部灵敏度分析

假设在，处可微，且强对偶性成立，那么有

，

最优Lagrange乘子就是关于约束扰动的局部灵敏度。

根据互补松弛性，如果，那么此约束不起作用，因此可以稍稍加强约束或者放松约束而不影响最优值。如果，即约束起作用，那么通过最优Lagrange乘子的第个分量可以知道此约束起作用的程度。如果较小，那么稍稍加强约束或者放松约束对于最优值没有大的影响；如果较小，那么即使稍稍加强约束或者放松约束对于最优值都影响很大。

#### 9．广义不等式

本节将Lagrange对偶理论扩展到具有广义不等式约束的问题，即



其中是正常锥。此时，并不假设问题是凸的。假设定义域非空。

* **Lagrange对偶**

对于上述问题，引入Lagrange乘子向量并定义相关的Lagrange函数如下



其中，。对偶函数的定义和原问题只有数值不等式的情形一样



因为Lagrange函数对于对偶变量是仿射的，且对偶函数是Lagrange函数的逐点下确界，所以对偶函数是凹的。对偶函数给出了原问题最优值的下界。

对于原问题值函数值不等式的情形，要求。类似地，此时对偶变量的非负约束被替换成如下条件



其中是的对偶锥。换言之，对应于广义不等式的Lagrange乘子必须是对偶非负的。

由对偶锥的定义可知，如果且对于原问题的任意可行点有。那么。因此



关于求下确界可以得到。

Lagrange对偶优化问题为



无论原问题是否是凸问题，弱对偶性总是成立，即，表示对偶问题的最优值。

Slater条件和强对偶性

广义Slater条件可以描述为：存在一点使得且。如果广义Slater条件成立，那么强对偶性成立（且可以达到对偶最优）。

* **最优性条件**

可以容易地将第5节内容推广至包含至广义不等式的问题。

互补松弛

广义不等式情形的互补松弛条件为



KKT条件



其中是函数在处的导数。

扰动及灵敏度分析

广义不等式情形下有



局部灵敏度分析同样成立：如果在处可微，则

