凸优化与组合优化复习

# 1.1矩阵求导

以分母布局为准。

如果结果是标量，对标量求转置还是它本身，注意合并。

这里和都是标量，两个值相等，因此：

# 1.2矩阵基础

## 1.2.1正定矩阵概念和性质

为实对称矩阵，如果正定，下面概念等价：

* 正定
* 的所有顺序主子式大于0
* 的所有特征值大于0
* 任意非零向量，有
* 存在一个列满秩实矩阵，使得

正定矩阵可逆。

如果半正定，上述概念中的大于0改为非负。存在一个行满秩实矩阵，使得。

半正定矩阵不可逆。

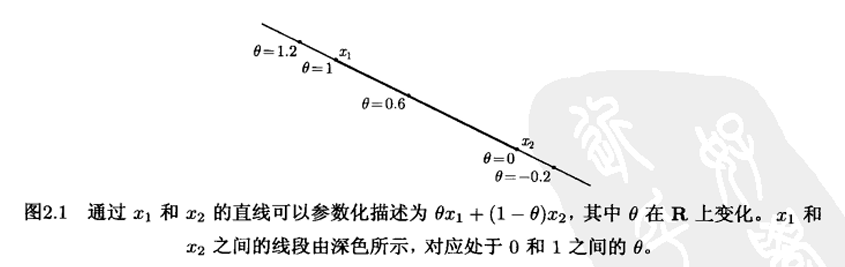
如果负定，下面概念等价：

* 负定
* 的偶数阶主子式大于0，奇数阶主子式小于0
* 的所有特征值小于0
* 任意非零向量，有
* 存在一个列满秩实矩阵，使得。

# 2.1凸集

## 2.1.1直线与线段

是基点（）和方向（由指向）乘以参数的和。



## 2.1.2仿射集

是仿射的等价于：对于任意，及，有。换言之，包含了中任意两点的系数之和为1的线性组合。

这个概念可以拓展到多个点的情况。如果，称为点的仿射组合。将仿射集拓展到多个点：如果是一个仿射集合，，并且，那么仍然在中。

如果是一个仿射集，并且，则集合

是一个子空间，即关于加法和数乘是封闭的。

证明对加法和数乘是封闭的：

1. 设，，，，则根据子空间定义，，。
2. 证明，的线性组合。考虑，由仿射集定义

则。

仿射集可以由子空间定义

定义的维数是子空间的维数。

线性方程组的解是仿射集。令方程组为，和是两个解，则对

因此也是的解。仿射集可以用的解集来表示。

仿射包：称集合中的点的所有仿射组合组成的集合为的仿射包，记为：

仿射包是包含的最小的仿射集合。

## 2.1.4凸集

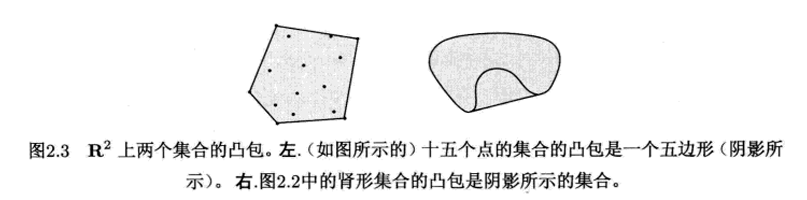
集合被称为凸集，如果中任意两点间的线段仍然在中，即对于任意，和满足的都有

称点为点的凸组合，其中并且。与仿射集合类似，一个集合是凸集等价于集合包含其中所有点的凸组合。显然，仿射集一定是凸集。（仿射集不限制和，因此凸集是仿射集的子集）



称集合中所有点的凸组合的集合为其凸包，记为：

凸包是包含的最小的凸集。



## 2.1.5锥

如果对于任意和都有，我们称集合是锥或者非负其次。如果集合是锥，并且是凸的，则称为凸锥，即对于任意和，都有

图表, 折线图

描述已自动生成

具有，形式的点称为的锥组合。如果均属于凸锥，那么，的每一个锥组合也在中。换言之，是凸锥的充要条件是它包含其元素的所有锥组合。

集合的锥包是中元素的所有锥组合的集合，即

它是包含的最小凸锥。

图片包含 图表

描述已自动生成

# 2.2重要例子

* 空集，单点集，全空间都是的仿射子集。
* 任意直线是仿射的。如果直线通过零点，则是子空间，因此也是凸锥。
* 线段是凸的，但不是仿射的（除非退化为一个点）。
* 一条射线是凸的，但不是仿射的。如果射线的基点是零点，则它是凸锥。
* 任意子空间是仿射的、凸锥。

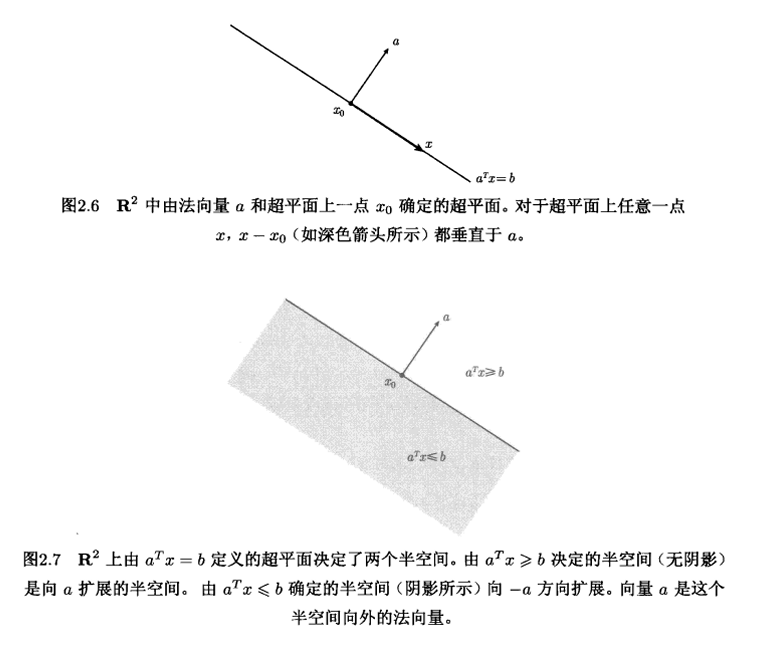
## 2.2.1超平面与半空间

超平面是具有下面形式的集合：

其中，，且。超平面是关于的非平凡线性方程的解空间（因此是一个仿射集合）。

一个超平面将划分为两个半空间。（闭的）半空间是具有下列形式的集合

即线性不等式的解空间，其中。半空间是凸的，但不是仿射的。



## 2.2.2球和椭球

中的球具有下面的形式

其中，，是2范数，是球心，标量是半径。由距离球心不超过的所有点组成。

是凸集，取球内一点和，则且。取和中间一点，则

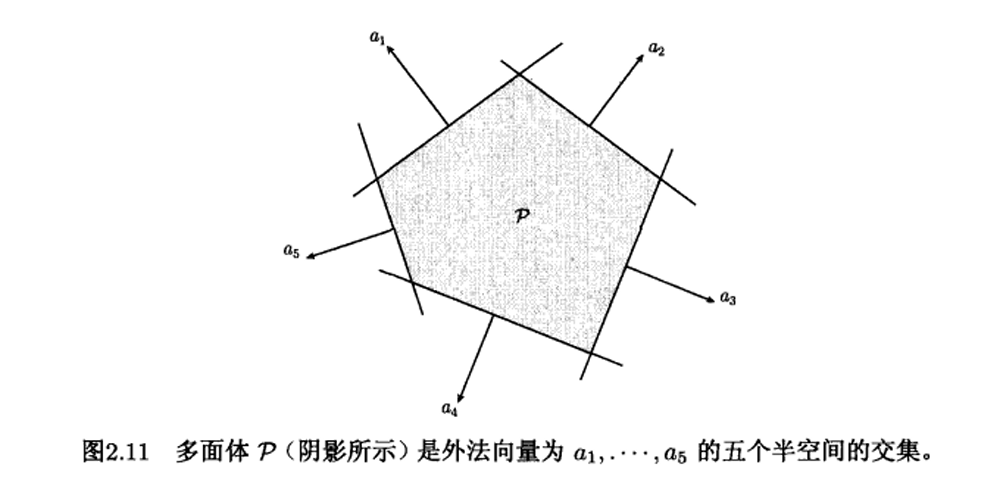
椭球是另一类凸集，具有下列形式：

其中，即是对称正定矩阵。的半轴长度由给出，是的特征值。

## 2.2.4多面体

多面体被定义为有限个线性等式和不等式的解集，

因此，多面体是有限个半空间和超平面的交集。仿射集合（子空间、超平面、直线）、射线、线段和半空间都是多面体。多面体是凸集。



可以使用

来表示定义，其中

此处表示上的向量不等式：表示每个分量。

单纯形是一类重要的多面体。设个点仿射独立，即线性独立，则这些点决定了一个单纯形。

其中**1**表示所有分量均为1的向量。这个单纯形的仿射维数为，因此也称为空间的维单纯形。

## 2.2.5半正定锥

用表示对称矩阵的集合，即

用表示对称半正定矩阵的集合，即

用表示对称正定矩阵的集合，即

这些符号与对应，表示非负实数，表示正实数。

半正定性和正定性：

是凸锥，对任意

是凸锥，对任意

不是凸锥。对任意，当时，，而不是。

# 3.1凸函数基本性质和例子

## 3.1.1定义

（定义1）

函数是凸的，如果是凸集，且对于任意和任意，有

从几何意义上看，上述不等式意味着点和之间的线段在函数的图像上方。

图示

描述已自动生成

称是严格凸的，如果上述不等式在且时严格成立。称是凹的，如果是凸的。称是严格凹的，如果是严格凸的。

所有的仿射函数是既凹又凸的。如果某个函数是既凹又凸的，那么它是仿射函数。

（定义2）

函数是凸的，如果是凸集，且对于，函数是凸函数。其定义域为。常用于化简证明。

## 3.1.2扩展值延伸

定义凸函数在定义域外的值为∞，从而将这个凸函数延伸至全空间，如果是凸函数，我们定义它的扩展值延伸如下：

类似地，可以通过定义凹函数在定义域外都为对其进行延伸。延伸后的函数保留的凸性或凹形。

## 3.1.3一阶条件

（定义3）

假设可微（即其梯度在开集内处处存在），则函数是凸函数的充要条件是是凸集且对于任意，下式成立

图示

描述已自动生成

一阶条件的证明：

首先证明一维的情况。

证明可微函数是凸函数的冲要条件是对于内的任意和，有

假设是凸函数，且。因为是凸集，因此对任意有，由的凸性可得

上式同除可得

令，可以得到一阶条件。

假设对内的任意和，函数满足一阶条件。选择任意，，令。应用一阶条件有：

将第一个不等式乘以，第二个不等式乘以，两者相加可得

从而说明函数是凸的。

再证明一般情况，即。设，考虑过这两点的直线上的函数，即函数，此函数对求导可得。

假设是凸函数，则函数是凸的，则有，即。

假设一阶条件对任意和均成立，因此若以及，我们有

即，说明函数是凸的。

## 3.1.4二阶条件

（定义4）

现在假设函数二阶可微，即对于开集内的任意一点，它的Hessian矩阵或者二阶导数存在，则是凸函数的要求是，其矩阵是半正定阵：即对于所有，有

对于上的函数，上式可以简化为一个简单的条件，此条件说明函数的导数是非减的。

二次函数，其定义域为，其表达式为

其中，，，。因为对于任意，，所以函数是凸的，当且仅当（是凹的当且仅当）。

## 3.1.5例子

* 指数函数。对任意，函数在上是凸的。
* 幂函数。当或时，在上是凸函数，当时，在上是凹函数。
* 绝对值幂函数。当时，函数在上是凸函数。
* 对数函数。函数在上是凹函数。
* 负熵。函数在其定义域上是凸函数。

范数。上的任意范数均为凸函数。范数应满足

证明范数是凸函数。

，

由定义1，范数是凸函数。

P范数：对于，，

无穷范数：

零范数：为中非零元素的数目

极大值函数：

证明极大值函数为凸函数。

，

指数和的对数：在上是凸函数。

文本, 信件

描述已自动生成

几何平均：在定义域上是凹函数。

文本, 信件

描述已自动生成

对数行列式：在定义域上是凹函数。

文本, 信件

描述已自动生成

## 3.1.6下水平集

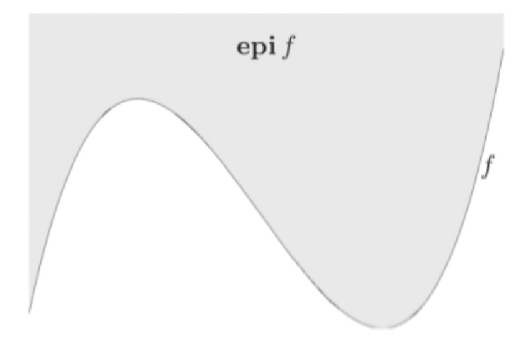
函数的-下水平集定义为

对于任意值，凸函数的下水平集仍然是凸集。证明可以由凸集的定义直接得到。反过来不一定正确：某个函数的下水平集都是凸集，这个函数可能不是凸函数。举例：。

## 3.1.7上境图

函数的图像定义为

函数的上境图定义为



一个函数是凸函数，当且仅当其上境图是凸集。

# 3.2保凸运算

## 3.2.1非负加权求和

如果函数是凸函数且，则函数也为凸函数。如果函数和都是凸函数，则它们的和也是凸函数。凸函数的集合本身是一个凸锥：凸函数的非负加权求和仍然是凸函数，即

是凸函数。

拓展到无限项的求和以及积分的情形。例如，如果固定，函数关于是凸函数，且对于任意，有，则函数

关于是凸函数。

## 3.2.2复合仿射映射

假设，，以及，定义为

其中。若函数是凸（凹）函数，则也是凸（凹）函数。

证明：

，

## 3.2.3逐点最大和逐点上确界

若和均为凸函数，则二者的逐点最大函数

其定义域为，仍然是凸函数。

这个性质可以拓展到若干个凸函数上。

证明：

，

最大r个分量之和函数是凸函数：

这是因为是个线性函数的逐点最大。

逐点最大的性质可以拓展到无限个凸函数的逐点上确界。如果对于任意，函数关于都是凸的，则函数

关于也是凸的。

对称矩阵的最大特征值。定义，，它是凸函数。

证明：

对取特征向量，有。则，。令，于是。。这是关于的一族线性函数的逐点上确界。

## 3.2.4复合函数

给定函数以及，定义复合函数为

给出复合函数规则：

* 是凸函数且非减，是凸函数，则是凸函数
* 是凸函数且非增，是凹函数，则是凸函数
* 是凸函数且非减，是凹函数，则是凹函数
* 是凸函数且非增，是凸函数，则是凹函数

表示的延伸，若点不在内，对其赋值或。要求在整个上非增或非减。

证明：

假设，。由于，我们有，且。因为是凸集，有，由的凸性有：

由，可得。假设是非减的，可以理解为其定义域在负方向上无限延伸。则上述不等式的左侧和右侧仍在定义域内，即，因此是凸集。

根据前提条件，是非减的，则

再由的凸性：

即：

复合定理得证。

简单的复合结论：

* 如果是凸函数，则是凸函数。
* 如果是凹函数且大于零，则是凹函数。
* 如果是凹函数且大于零，则是凸函数。
* 如果是凸函数且不小于零，，则是凸函数。
* 如果是凸函数，则在上是凸函数。

# 3.3共轭函数

设函数，定义函数为

此函数称为的共轭函数。在二维空间下，相当于给定斜率，确定使得线性函数和之间得到最大差值。

图示, 折线图

描述已自动生成

是凸函数，因为它是一系列的凸函数的逐点上确界。

* 若可微，则对应的必是的一点
* 函数的共轭一定是凸函数

常见函数的共轭函数。

* 仿射函数。。。
* 负对数函数。，定义域为。当时，函数无上界，当时，在处函数达到最大值。因此，
* 严格凸的二次函数。。令，。

# 3.4拟凸函数

## 3.4.1定义及例子

（定义1）

函数称为拟凸函数（单峰函数），如果其定义域及所有下水平集

其中，，都是凸集。函数是拟凹函数，如果是拟凸函数，即每个上水平集是凸集。若某函数既是拟凸函数又是拟凹函数。称其为拟线性函数，即其定义域和所有水平集都是凸集。

凸函数具有凸的下水平集，所以它是拟凸函数。但拟凸函数不一定是凸函数。

（性质，定义2）

是拟凸函数当且仅当是凸集，且对于任意及，有

即线段中任意一点的函数值不超过端点函数值较大的那个。

图示

描述已自动生成

例子：

* 向量的长度。，中最后一个非零元的位置。

令，得到的集合是子空间，所以是拟凸函数。

* 向量的零范数。非零元素的个数。

## 3.4.3可微拟凸函数

（定义3，一阶条件）

设函数可微，则函数是拟凸的充要条件是，是凸集，且对于任意有

与凸函数定义不同，凸函数中，是全局最小点。而拟凸函数中这样的论断不成立。

（定义4，二阶条件）

假设函数二次可微。如果函数是拟凸函数，则对任意和有

即在斜率为0的点，二阶导数是非负的。

# 3.5对数-凹函数和对数-凸函数

称是对数凹，如果对所有有且是凹函数。

称是对数凸，如果对所有有且是凸函数。

若：

* 若凸，不一定对数凸。
* 若对数凸，则一定凸。
* 若对数凹，则不一定凹。
* 若凹，一定对数凹。

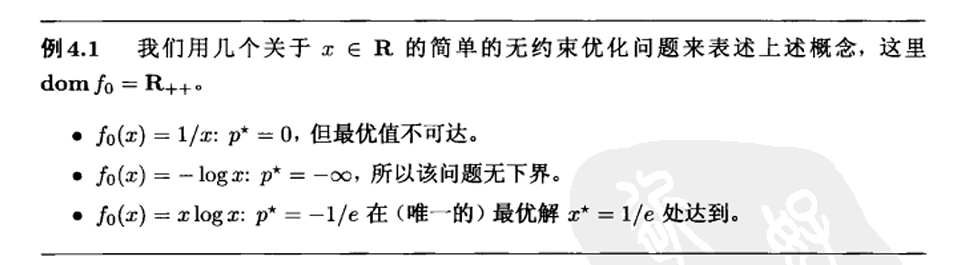
# 4.1优化问题

## 4.1.1基本术语

定义域。

令，称为可行解集。且满足约束称为可行解。

令为问题最优解。允许。



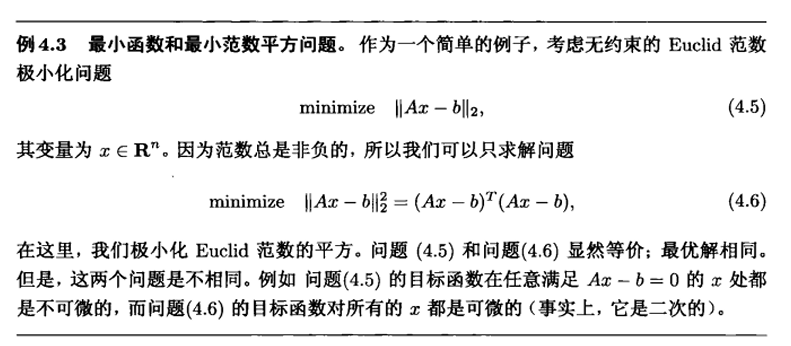
## 4.1.3等价问题

非正式的定义：如果从一个问题的解，很容易得到另一个问题的解，并且反之亦然，称两个问题等价。

变量代换：令，则原来关于的问题与变换后关于的问题等价。需要说明：如果解决了原问题，那么也解决了变换后的问题；如果解决了变换后的问题，那么解决了原问题。

目标函数变换：如果单增，那么优化和等价。

约束函数变换：如果，则约束与等价。如果，则约束与等价。



消除不等式约束：引入松弛变量。

消除等式约束：将等式约束显式代入。

以线性等式约束为例：

问题有线性等式约束，可以求出，其中是零空间的矩阵。

文本, 信件

描述已自动生成

# 4.2凸优化

## 4.2.1标准形式的凸优化问题

三个附加要求：

* 目标函数是凸的。
* 不等式约束函数是凸的。
* 等式约束函数是仿射的。

凸优化问题的可行集是凸集，因为凸优化问题的定义域是凸集，而可行集是与定义域的交集。

如果是拟凸的而非凸的，那么它是拟凸优化问题。

文本, 信件

描述已自动生成

## 4.2.2局部最优解与全局最优解

凸优化问题的任意局部最优解也是全局最优解。

拟凸优化问题没有这一性质。

## 4.2.3可微目标函数的最优性准则

令表示凸优化问题的可行集。那么是最优解，当且仅当且满足

证明略。

由该条件可推导出一些简单的例子：

无约束问题。是最优解充要条件为

只含等式约束的问题。

文本, 信件

描述已自动生成

是最优解充要条件为。

只含有非负约束的问题。

文本, 信件

描述已自动生成

是最优解充要条件为，这说明要么且，要么且，称为互补性条件。

# 4.3线性规划问题

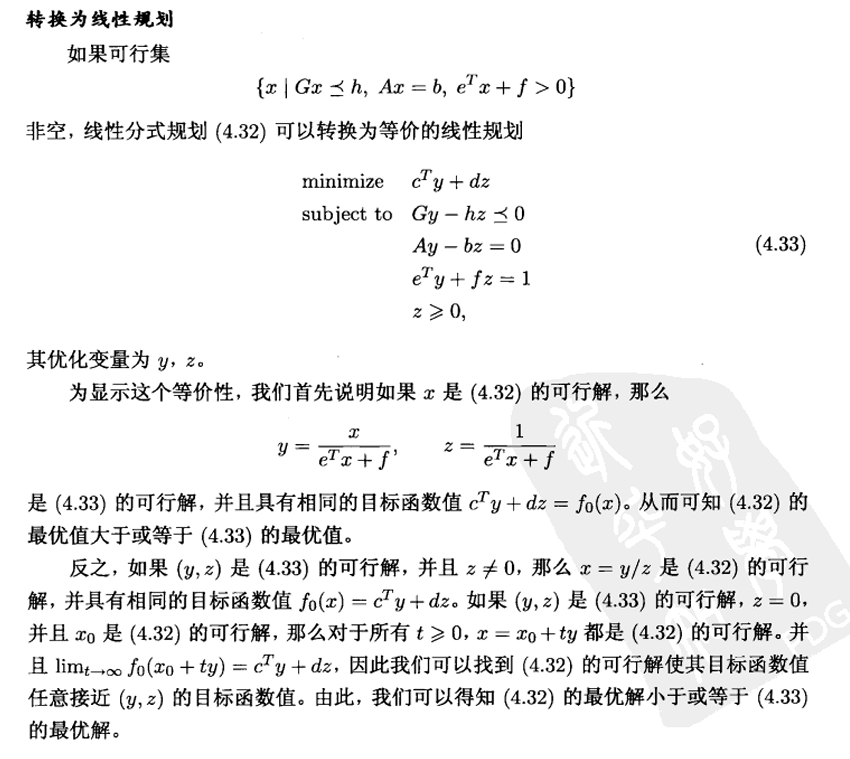
线性规划的标准形式：

将线性规划转换为标准形式：

1. 引入松弛变量消除不等式约束：
2. 将无约束的变量表示为两个非负变量和的差：

## 4.3.2线性分式规划

这个目标函数是拟凸的，这个问题是拟凸优化问题。



# 4.4二次优化问题

当凸优化问题的目标函数是二次型并且约束为仿射时，该问题称为二次规划。即

其中，

如果不等式约束是二次型，即

称之为二次约束二次规划（QCQP），其中。

# 5.1拉格朗日对偶函数

## 5.1.1拉格朗日函数

## 5.1.2拉格朗日对偶函数

这是关于取得的最小值。是问题的定义域。，。如果拉格朗日函数关于无下界，则对偶函数取值为。对偶函数是关于和的仿射函数的逐点下确界，所以对偶函数一定是凹函数。注意，原问题不一定是凸优化问题。

## 5.1.3最优值下界

对偶函数在和任意时，下式成立：

这是因为

称以及的是对偶可行的。否则任意的使意义不大。

## 5.1.5例子

* 线性方程组的最小二乘解

写出拉格朗日函数：

写出对偶函数：

是的二次凸函数，无约束下最优性条件为：

因此，代入有，这是原问题最优值的一个下界。

* 标准形式的线性规划

写出拉格朗日函数：

写出对偶函数：

是一个线性函数，其下确界可取。于是

只有和时，下界才有意义。

* 双向划分问题

其中，是对称阵。

写出拉格朗日函数：

写出对偶函数：

是一个二次函数，若，则其下确界为0。否则，其下确界为。因此

# 5.2拉格朗日对偶问题

问题：从拉格朗日函数能够得到的最好下界是什么？

这是一个凸优化问题，因为是凹函数，且约束集合是凸集。记为对偶问题的最优值，则。

## 5.2.1显式表达对偶约束

标准形式的线性规划问题的对偶问题的对偶问题与原问题等价

写出对偶函数（见上述例子）

写出对偶问题

我们将约束显式化得到等价问题

进一步地，将代入约束，将目标改为最小化，最后将改为有

我们对这个问题再做一次对偶

拉格朗日函数：

对偶函数：

对偶问题：

这个问题与最开始的问题等价。

因此，线性规划问题的对偶问题的对偶问题是原问题。。

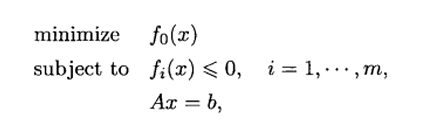
## 5.2.2弱对偶性

是一个重要不等式。即使当和为无限时，不等式仍然成立。这就是说时，一定有，即对偶问题不可行。反过来，时，一定有，即原问题不可行。

## 5.2.3强对偶性

成立时，称强对偶性成立，这说明从拉格朗日对偶函数得到的最好下界是紧的。凸问题的强对偶性通常（但不总是）成立。

考虑凸问题：



相对内部：记集合的相对内部为，

是半径为，中心为并由任意范数定义的球。是仿射包。可以简单理解为D的内部点集。

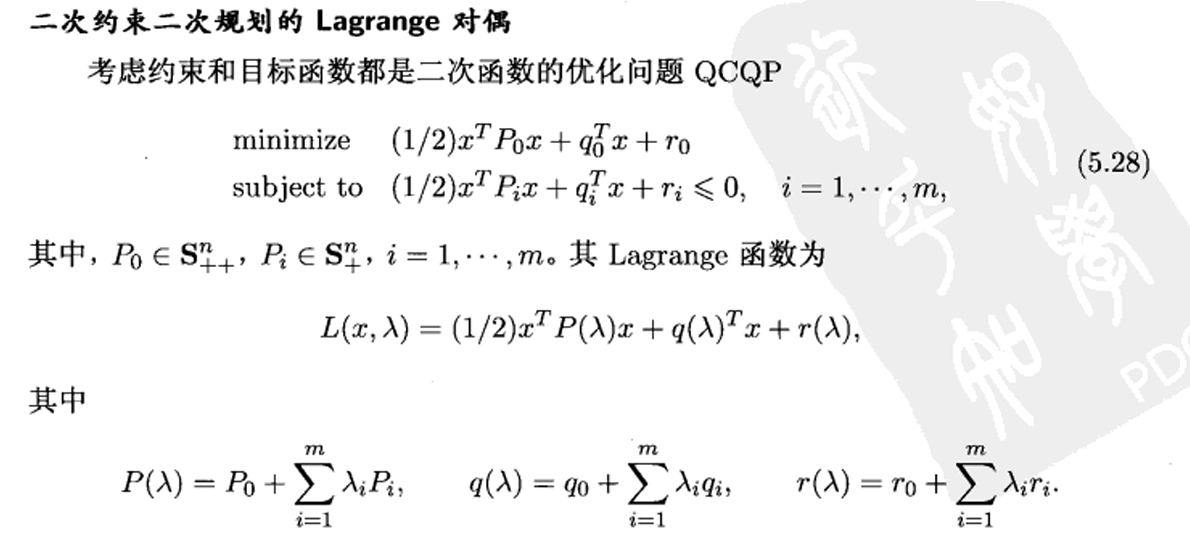
Slater条件：存在一点使得

当Slater条件成立且**原问题是凸问题**时，强对偶性成立。这是一个充分不必要条件。

弱化的Slater条件：

仿射不等式不需要严格成立，其他不等式条件严格成立。这意味着若所有约束等式都是线性等式和线性不等式且是开集时，弱化的Slater条件总成立。

二次约束的二次规划问题：



直接考虑的情况，此时有，

此时对偶问题为：

文本, 信件

描述已自动生成

根据Slater条件，当二次不等式约束严格成立时，即

强对偶性成立。

我们考虑。此时问题简化为

此时的Slater条件是严格成立，由于，这样的点不存在，于是Slater条件不成立。但此时，由于且，因此。强对偶性成立。

# 5.3几何解释

略。

# 5.4鞍点解释

## 5.4.1强弱对偶性的极大极小描述

假设没有等式约束（事实上可以很容易拓展到等式约束的情况）

假设不可行，即存在某些使得，选择，其他，则。如果可行，则每个，此时选择每个，则。我们可以说

根据对偶函数的定义：

则

这实际上是对同一个函数的极大极小（极小极大）过程，对任意函数都有极大极小不等式成立：

## 5.4.2鞍点解释

考虑任意函数定义在上，称使得

成立的点为的鞍点，等价地可写为

因此，（鞍点定理）如果和分别是原问题和对偶问题的最优点，且强对偶性成立，则它们是拉格朗日函数的一个鞍点，反过来同样成立。

证明：

先假设为的鞍点，那么使得

成立，从而。

再假设为最优解且强对偶性成立，则可行，即

由强对偶性：

（第一个等式是强对偶性，第二个等式是g的定义，第三个等式是L的定义，第四个不等式是极小化定义，第五个不等式是因为且）

因此

又因为

所以

# 5.5最优性条件

## 5.5.2互补松弛性

鞍点定理的证明暗含了一个重要的信息：

因此，

而，因此每一项，这称为互补松弛条件，可以写成

这说明如果最优解x使某个或某些不等式约束达到0，则意味着x位于这个或这些约束的边界，这样的约束称为active constraint。【最优解没有取在不等式约束边界，不等式约束无效】，称这种约束为inactive constraint，因此可以直接“扔掉”，令。

## 5.5.3KKT条件

鞍点定理的证明暗含了另一个重要的信息：

即：在得到和后，我们对拉格朗日函数做极小化，能得到原问题的最优解。假设目标函数和约束函数都可微（但不假设它们是凸函数），因此函数在处的导数必须为零（否则，总能在附近沿负梯度方向以小步长找到一个更优解），即

因此有：

——互补松弛条件

——稳定性条件

上式称为Karush-Kuhn-Tucker (KKT) 条件。如果强对偶性成立，那么任一对原问题最优解和对偶问题最优解必须满足KKT条件。这是一个必要条件。

如果原问题是凸问题，各个函数都可微，强对偶性成立，则KKT条件为充要条件。

证明：

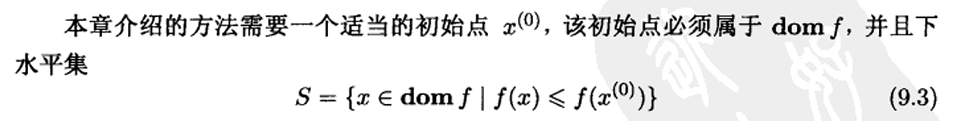
设满足KKT条件。

是凸函数，再由稳定性条件知为的最小值点，则。

又因为

故为的鞍点，由鞍点定理知，为原问题最优解。

# 9.1无约束优化问题



## 9.1.2强凸性

文本, 信件

描述已自动生成

注意是单位矩阵。

利用强凸性说明与距离的上界：

式9.8右侧可以看做是关于的凸函数，由凸函数的最优性条件：

解得，代入9.8右侧有：

因为是任取的，因此最优值。进一步地，因为，。

这说明当梯度足够小时，任意离最优值足够接近，可以当作次优解。

利用强凸性说明与的距离上界。

文本, 信件

描述已自动生成

因为是闭集，一定有。通常可取矩阵的的最大特征值，而可取的最小特征值。称比值是的条件数。又定义凸集的条件数，只需知道下水平集的条件数上界为。

# 9.2下降方法

通用下降方法：

1 给定初始点

2 重复进行下面的步骤

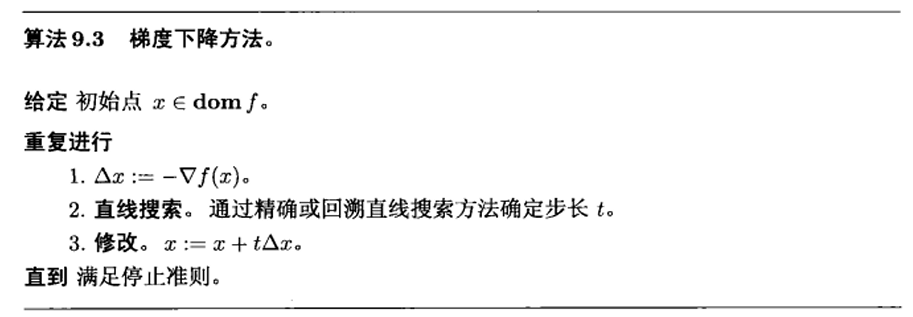
3 确定下降方向

4 直线搜索。选择步长。

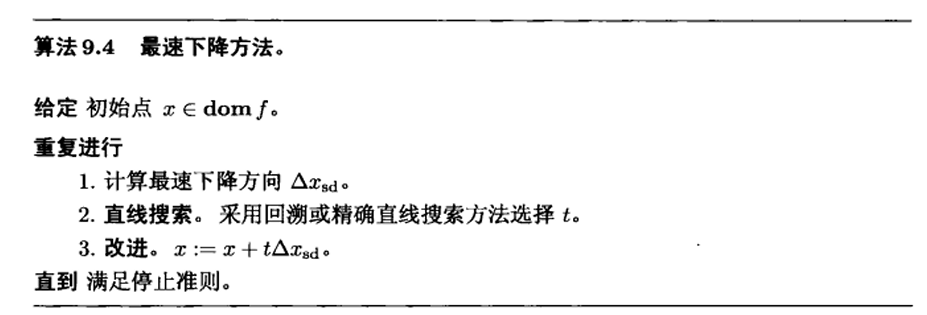
5 更新：。

6 直到达到某个收敛标准。

# 9.3梯度下降方法

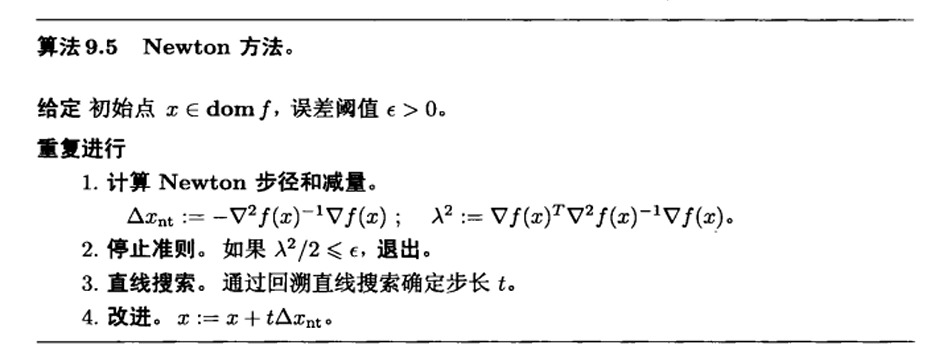


# 9.4最速下降法



其中，，称为规范化的最速下降方向。来自一阶泰勒展开

# 9.5牛顿法



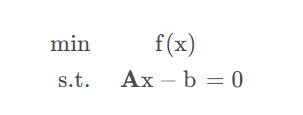
文本, 信件

描述已自动生成

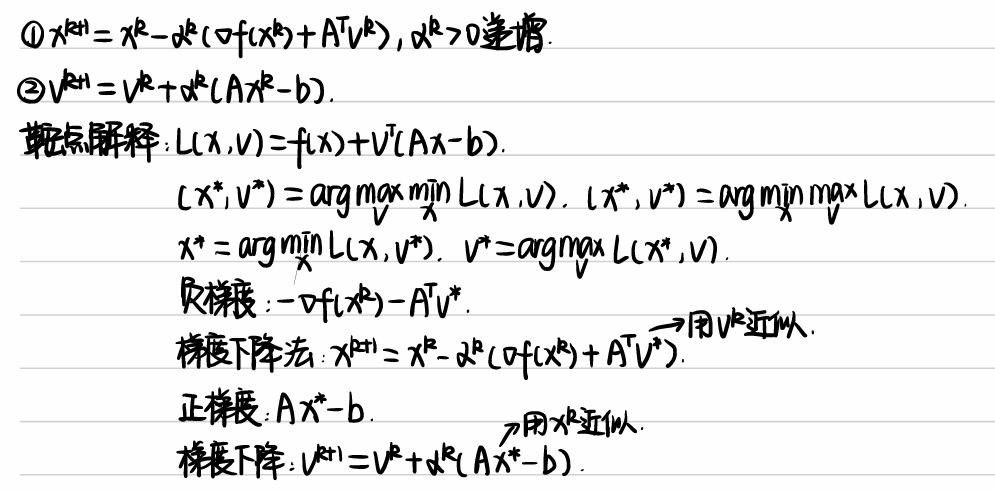
纯牛顿法不考虑，且停止准则放在最后。

# 拉格朗日法

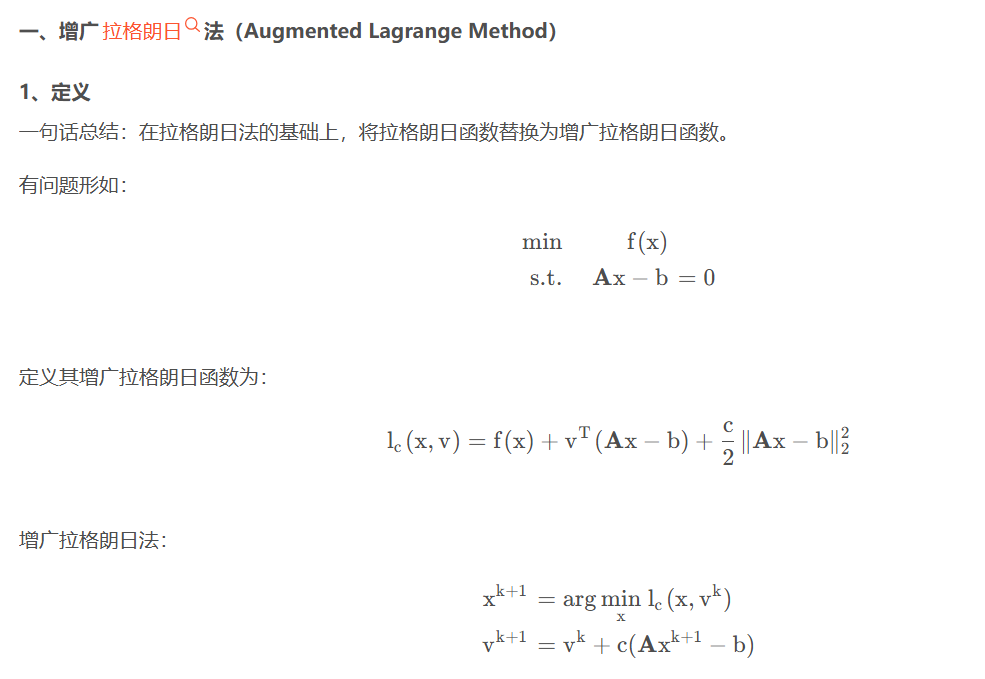
有问题形如：



其拉格朗日函数。



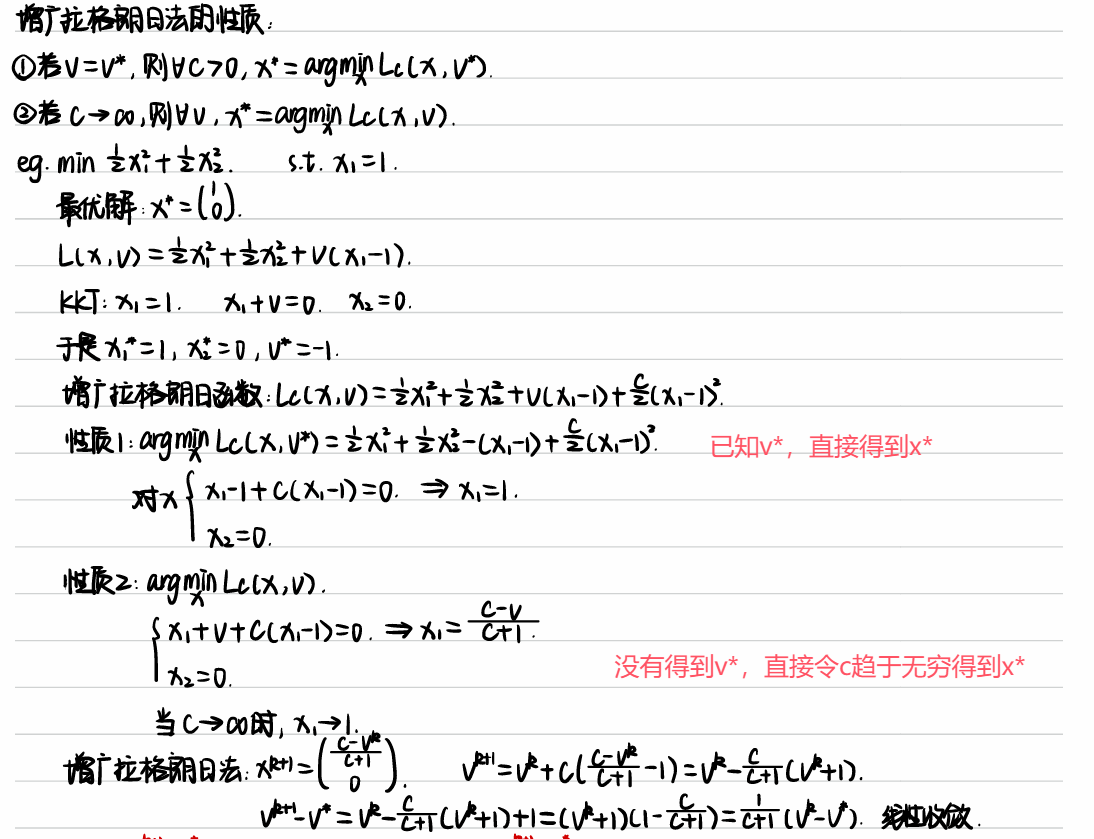
# 增广拉格朗日法



增广拉格朗日法相当于用拉格朗日法解决

对于这个新问题：

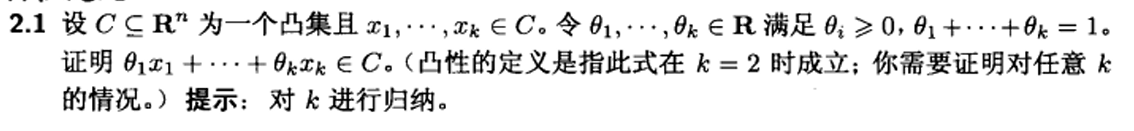
1. 新问题的最优解和原问题的最优解相同。这是容易证明的。
2. 新问题的对偶问题的最优解和原问题的对偶问题的最优解一样



# 可能出现的题型

## 凸集证明

证明凸性。

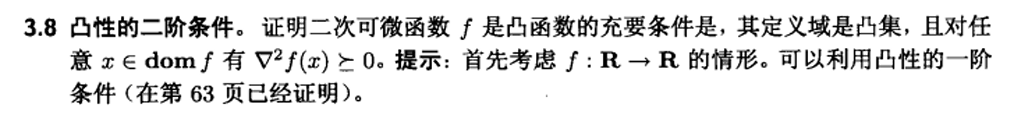


解：n=2时显然成立；假设n=k-1时成立；当n=k时，对，总能拆成，只需证明即可。而，满足n=k-1的归纳假设，因此。所以n=k时，。

证明某个例子是凸集。（都用定义证明）

## 证明某个函数是凸函数/拟凸函数（凹函数/拟凹函数）

证明二阶条件。



证明某个例子是凸函数/拟凸函数。

文本, 信件

描述已自动生成

1. ，凸函数。
2. Hessian矩阵：，既不是正定阵也不是负定阵，所以它既不是凸函数也不是凹函数。其上水平集是凸集，它是拟凹函数。
3. Hessian矩阵：，凸函数。
4. Hessian矩阵：，既不是正定阵也不是负定阵。
5. Hessian矩阵：，凸函数。

以及一些常见例子

指数和的对数：在上是凸函数。

几何平均：在定义域上是凹函数。

文本

描述已自动生成

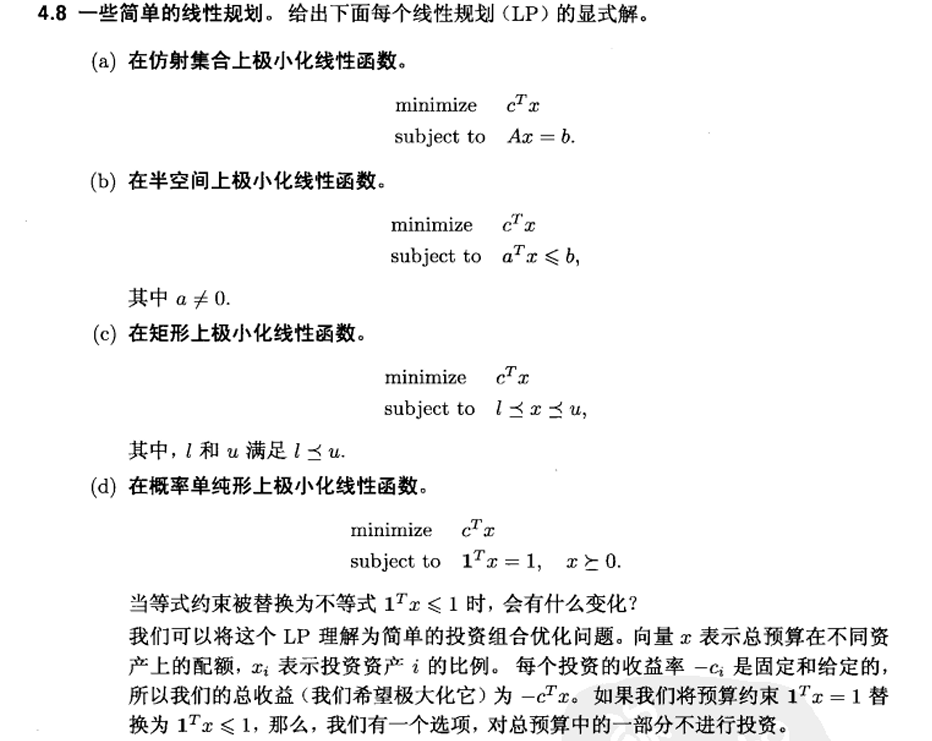
文本, 信件

描述已自动生成

1. 是凸的，是凹的，是凸函数且单调减，所以是凸的。

## 解决常见的凸优化问题

* 1. 线性规划



单纯形法

图片包含 文本

描述已自动生成

验证某解为最优解：最优性条件/强对偶性对偶最优

* 1. 二次规划

图示

描述已自动生成

文本, 信件

描述已自动生成

## 判断优化问题之间是否等价

思路：从原问题的可行解，构造出目标问题的可行解。再从目标问题的可行解，构造出原问题的可行解。

## 对偶问题与最优性条件

鞍点定理证明

可微凸问题下，KKT条件充要性证明

文本, 信件

描述已自动生成

1. 可行集为，最优值为，最优解为。
2. 拉格朗日函数

对偶函数

1. 对偶问题

文本, 信件

描述已自动生成

文本

描述已自动生成

文本, 信件

描述已自动生成

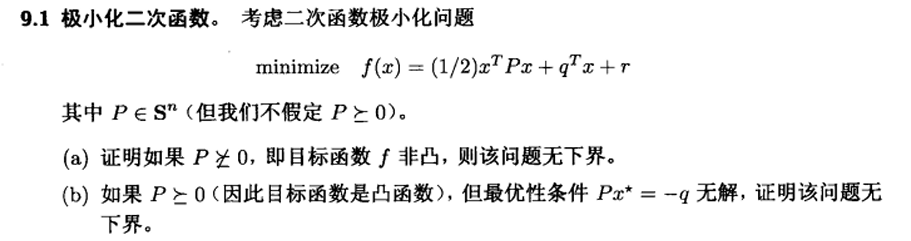
文本

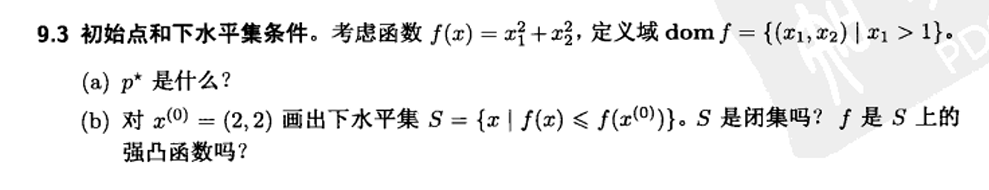
描述已自动生成

图片包含 文本

描述已自动生成

## 常见的优化算法







文本, 信件

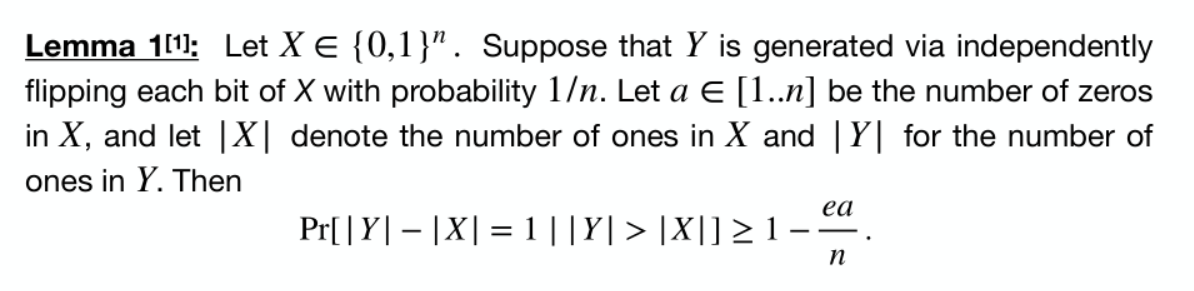
描述已自动生成



文本

描述已自动生成

## 智能优化算法



文本, 信件

描述已自动生成

文本, 信件

描述已自动生成

