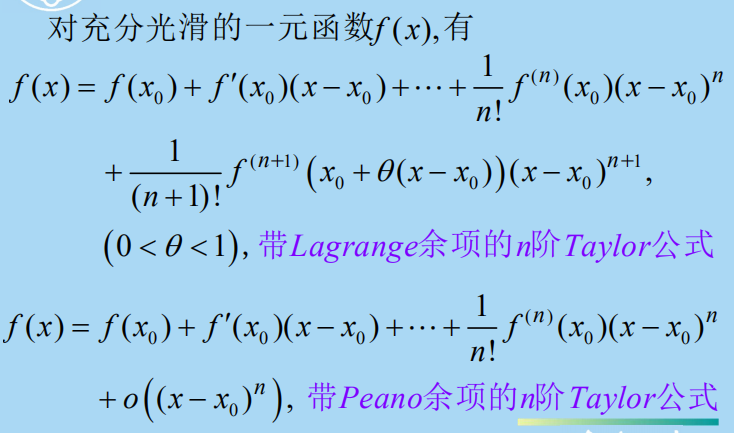
多元泰勒公式

4月14日14:44

1. 基本概念：

A. 类比于一元函数的泰勒公式：



我们想对多元函数来一次用n阶导数乘以变化量的n次方的展开；

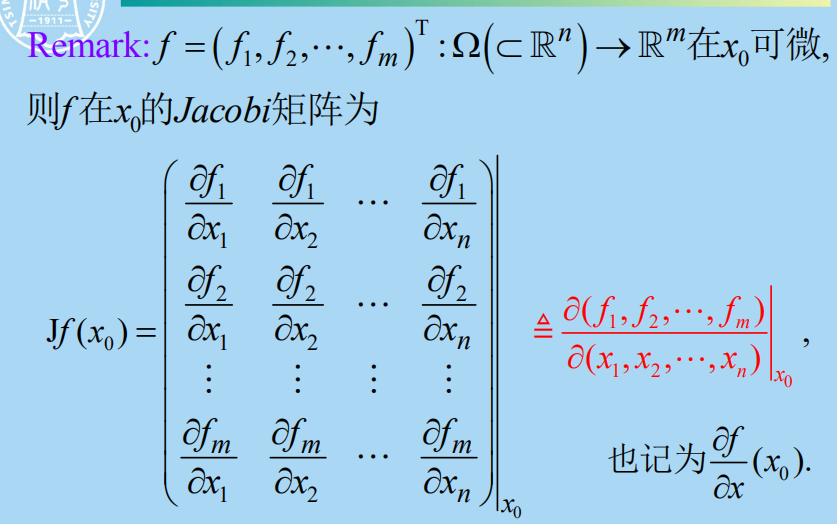
什么是多元函数的一阶导？很自然的想到一阶偏导数。

什么是多元函数的二阶导？很自然地想到二阶偏导数。

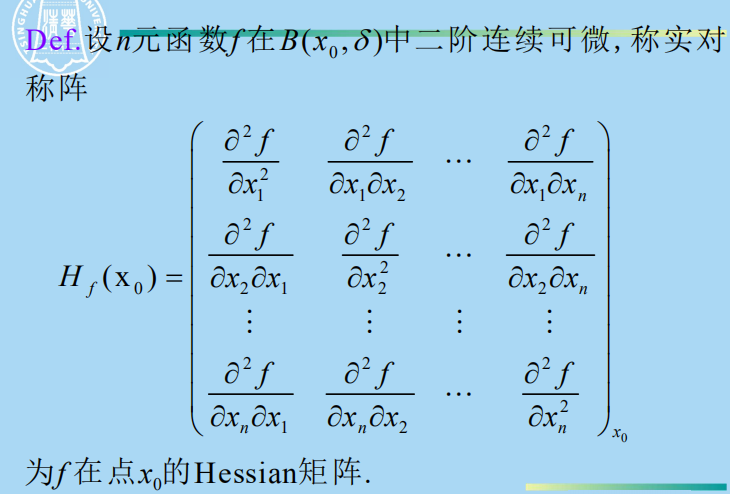
1. 建立定义：

一阶偏导函数，我们已经定义过，也就是Jacobi矩阵：

并且注意到，这里f是个一维的向量值函数，故而jacobi只有一行。



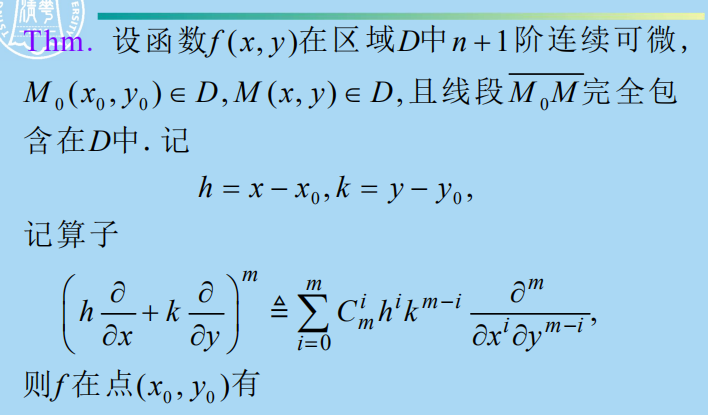
二阶偏导数，定义为Hessian矩阵：



注意到f二阶连续可微，故而Hessian矩阵是实对称的。

其实记住二阶黑塞矩阵的形式，其他的都好记，看准对角线。

完成了这些定义之后再来定义Taylor算子：



这个taylor算子非常不说人话，下面把他翻译为人类语言：



就是杨辉三角的展开，注意到k的阶次和对y求的偏导必然相等。

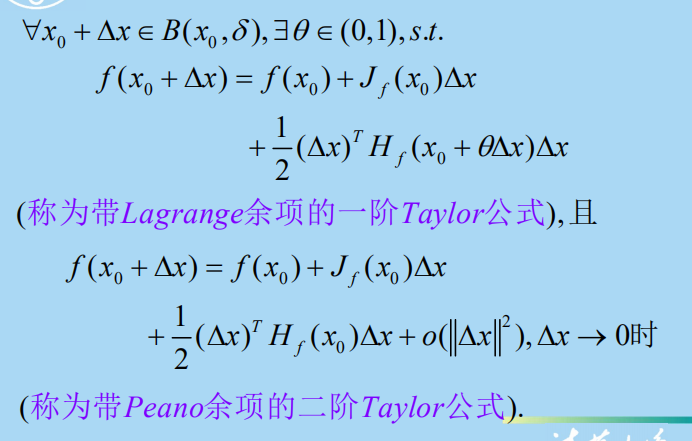
类比下也可以定义出三元函数的展开。

以及，Taylor算子的运用非常有限，仅仅用于算拉格朗日余项；

比如计算：的带有二阶Lagarange余项的泰勒 公式；带有二阶Lagarange余项的泰勒公式：余项是二阶的，多 项式是一阶的。只有拉格朗日余项需要你把具体的导数算出来。

以上完成了Jacobi、Hessian、Taylor算子的定义；

接下来定义Taylor公式：



带拉格朗日余项的n阶Taylor公式是指展开到n阶，余项是n+1阶；

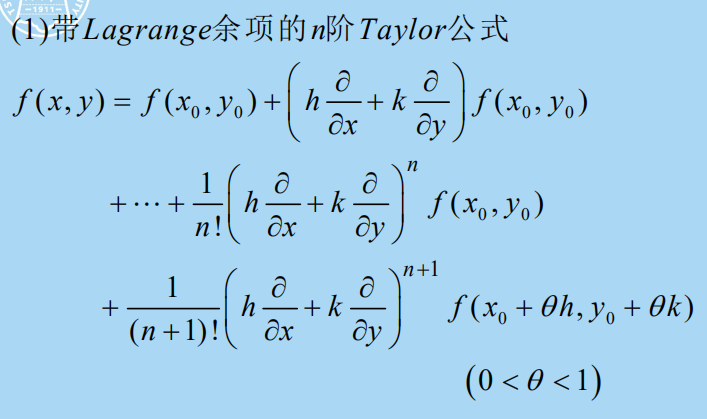
带Peano余项的n阶Taylor公式是指展开到n阶，余项是n阶；

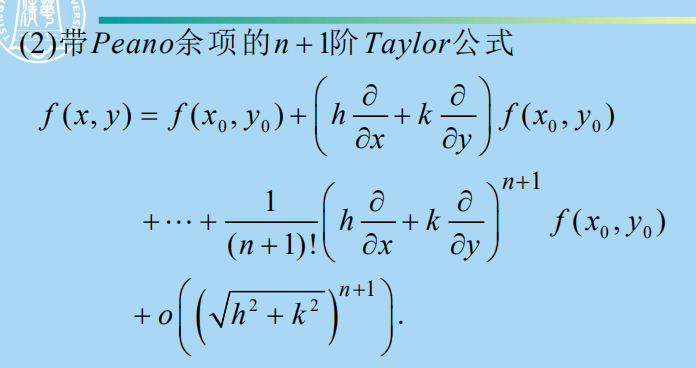
这个其实很好理解，因为peano余项虽然是O(x的n次方)实际上里 面是至少是x的n+1次方；

故而，n阶泰勒公式都是指泰勒多项式为n阶；但是带有n阶某某余项的泰勒公式，对于拉格朗日而言是展开到n-1阶，对Peano来说是展开 到n阶；（当然老师说了，不会出现第二种情况）

以及，泰勒公式是带有余项的；泰勒多项式是不带余项的；

同样，别的写法定义高阶二元Taylor公式：





注意到前面定义Taylor算子的时候没有带上阶乘，一定不要忘了！ 特别是在用Taylor算子求拉格朗日余项的时候；

二、计算方法：

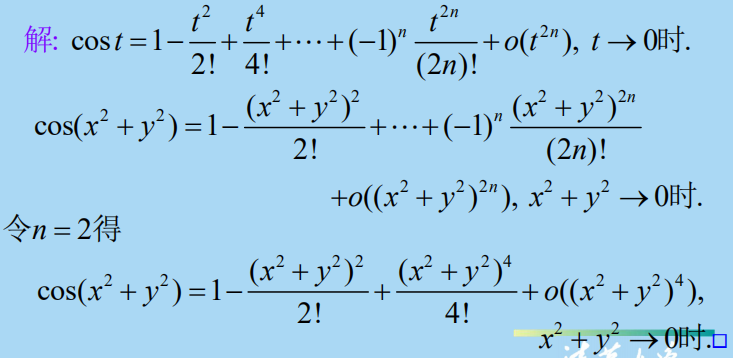
A.能够完全类比为一元泰勒：

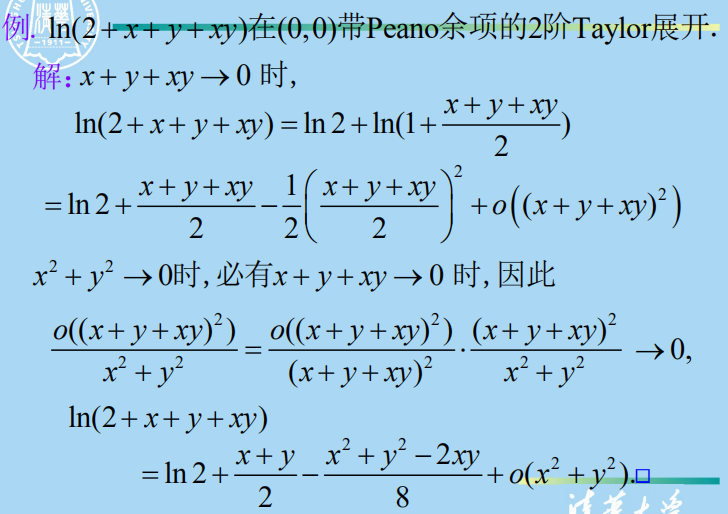
适用于Peano与纯粹的泰勒多项式，不需要通过高阶导表达拉格朗日余项，或者由于偏导函数的对称性，拉格朗日余项很好表达；



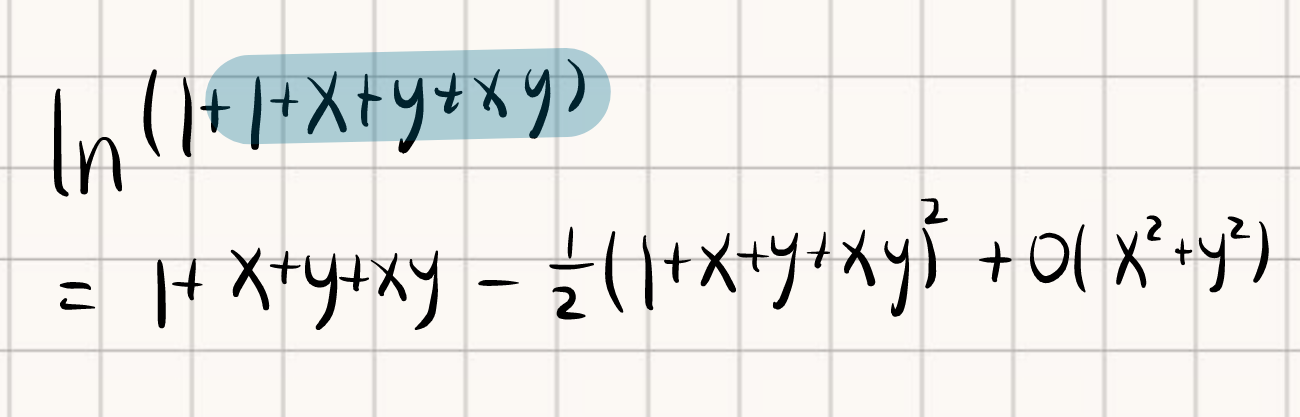
直接对cost展开到四阶（因为带入到t的四次方后就是八阶），然后保留四阶Peano余项；

注意到：n阶泰勒公式对于x和y的次数最高就是n。余项必然是O（x\_n）或者n+1阶拉格朗日；





这个例子非常好！



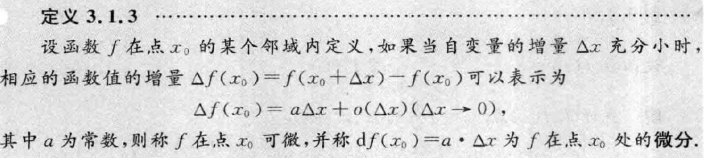
有奖竞猜，图中一共有多少错误？（数对了奖励亲亲一个，呜）

//1+x+y+xy不是无穷小量

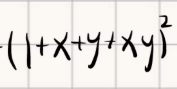
//二次项里面混进了四次项

首先，本质问题——

Peano余项的定义：



你必须要在△x趋于0时才能够定义高阶无穷小；如果你的 △x本身不趋于0，比如上面那个你△x=1+x+y+xy，本来就没 法定义高阶无穷小，所以一定是错的；

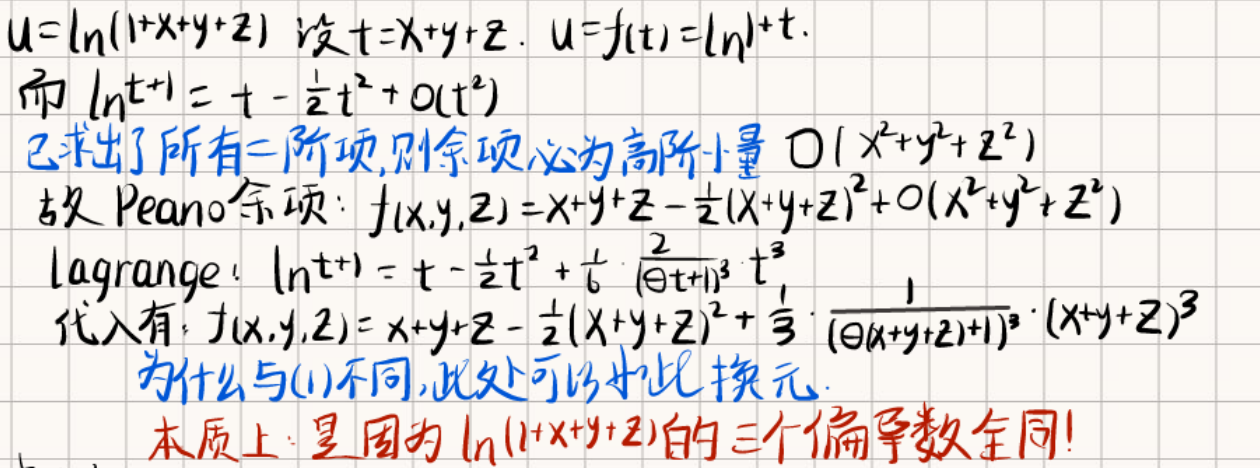
其次：n阶泰勒多项式，最高次一定是n次；xy是二次的，如 果你保留了，那最高次是四次的，并不是展开到2 阶，所以要删去二阶以上的项。

1. 其实最后写余项的时候，因为你前面已经出现了x和y的平方，把高阶项舍弃后，余项可以放心大胆地写成（当然，这是对(0,0)展开的余项）

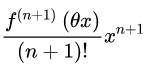


对他的展开，首先是在原点处的Peano，那么套用一元展开（一元泰勒也是要抽背的）。

接下来是原点处的Lagrange，注意到偏导数的对称性，所以也可以类比一元展开。



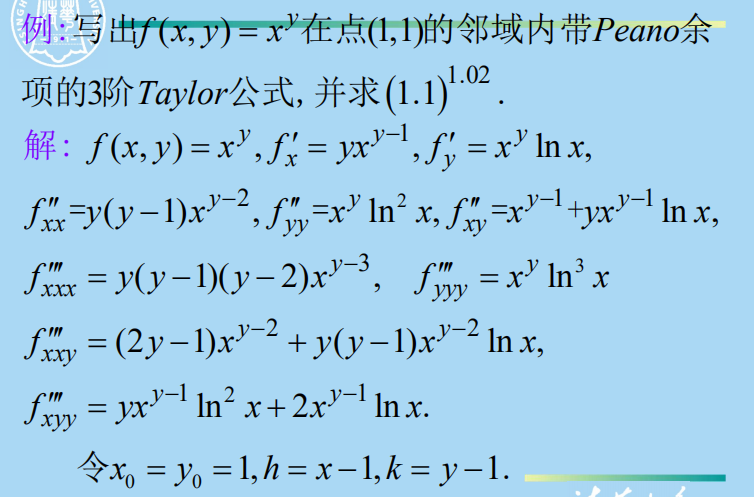
我发现我连都没理解到，

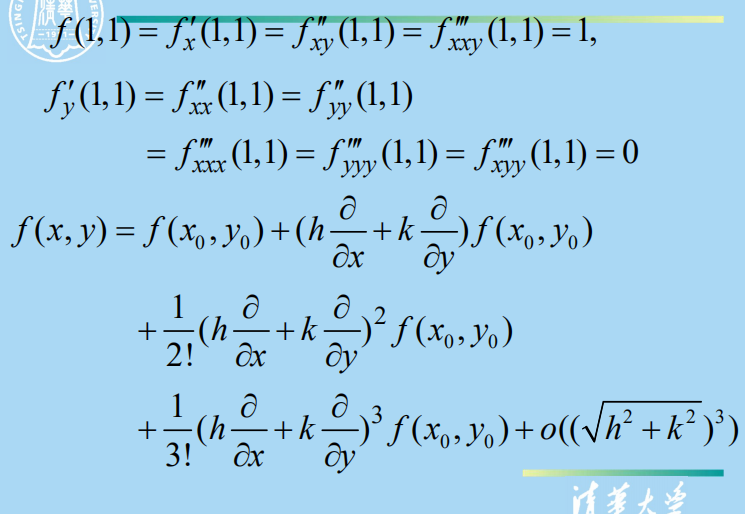


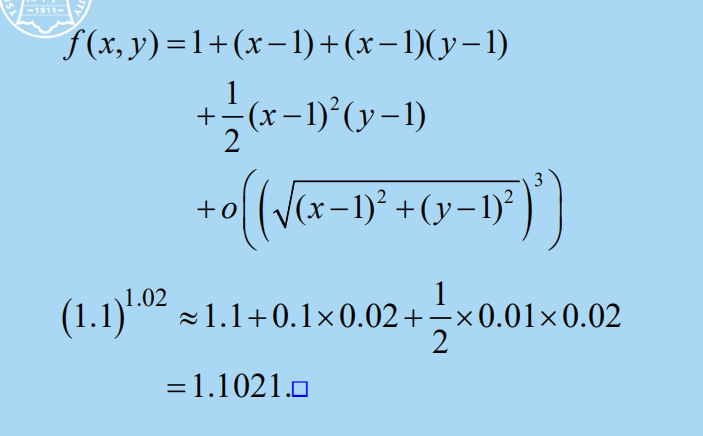
注意到，最后的1/3不是因为我们背的麦克罗林公式，而是碰巧三阶导有一个2从而抵消了3！里的一个2；也就是说，lagrange余项一定要你老老实实把高阶导求出来；

1. 无法类比为一元函数的Taylor公式：

//老实求偏导，不要试图进行奇怪的一元泰勒展开



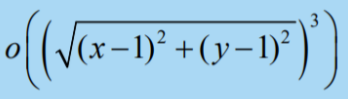




当然，其实这里求具体点偏导数的值并不是很优化，优化的方法是在最后一次求导的时候无关变量带入常数。

老实求出偏导，然后带入Xo给出偏导数；

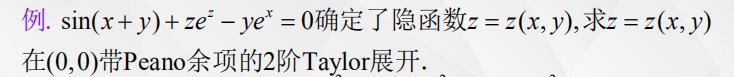


注意到，这一部很关键，把h和k表达出来简化计算，同时也避免了你的余项盲目写成，因为余项实际上是：！！！

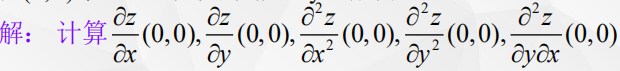
另外，求高阶偏导数的时候，所有没有被微分的变量都可以带入常数值， 但是被微分了的变量一定不可以！所以这个方法用于泰勒比较局限；

但是：在求最后一步时，比如已经有了二阶偏导数且没有带入过常数，这个时候带入常数当然能够很方便地算三阶偏导数。

C.必考题型，隐函数泰勒展开：

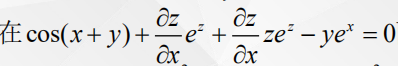
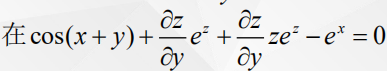


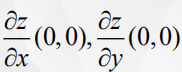
理清楚目标，既然无法套用一元函数的公式，那就是要求：

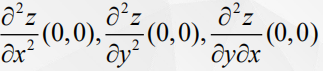


这些东西怎么来，无非就是对F(x,y,z)=whatever两边求导，注意到你在求偏导的过程中，whatever不影响偏导本身，但是影响你最后的那个点；（这其实还是超平面的观点，whatever直接决定了具体的点，F（x,y,z）决定了点的梯度，偏导etc）

求完偏导后得到这俩式子，注意到，还是默念，我在对谁求偏导？

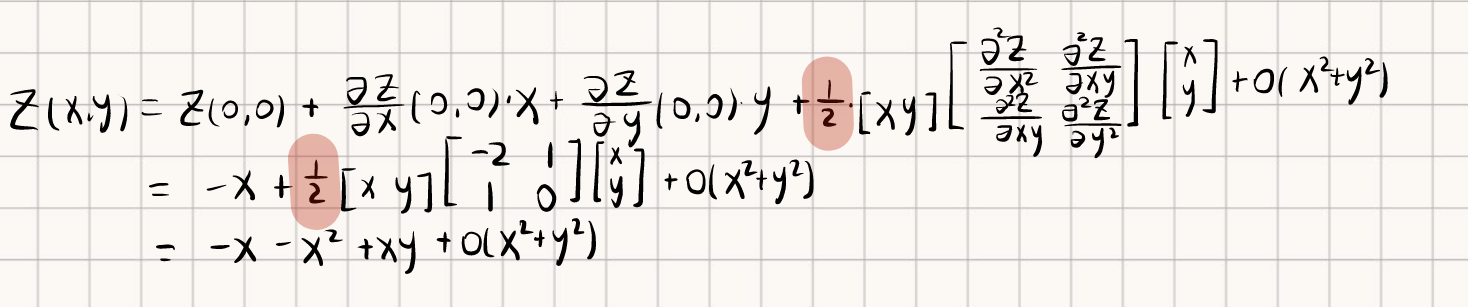


注意到，不要把这俩式子化简，直接带入Z=(0,0)=0进去解出俩一阶偏导；，再对上面两个式子求导，解出



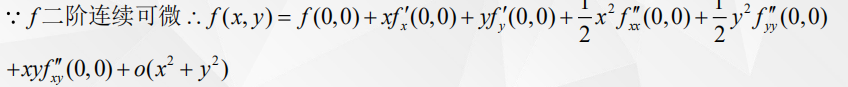
如法炮制，不要化简。

最后求出了五个需要用到的偏导数，再带入公式即可：

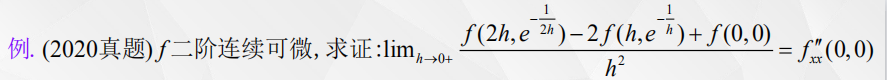


注意不要忘了公式里的阶乘项！！！

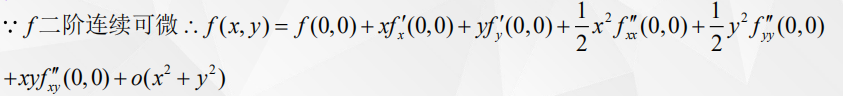
实际上我们可以直接给出：



D.泰勒展开的证明题：

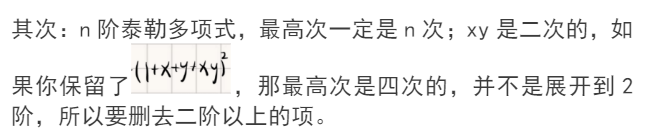


这个题注意到如果用洛必达，其实很难，因为你的目标，二阶偏导数很难构造，所以我们选择用泰勒，并且二阶连续可微看上去就是为泰勒创建的，直接展开到二阶；





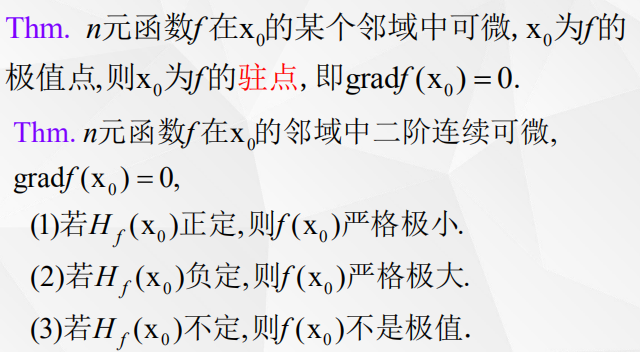
这一步非常神奇，请问和y有关的偏导去哪儿了？

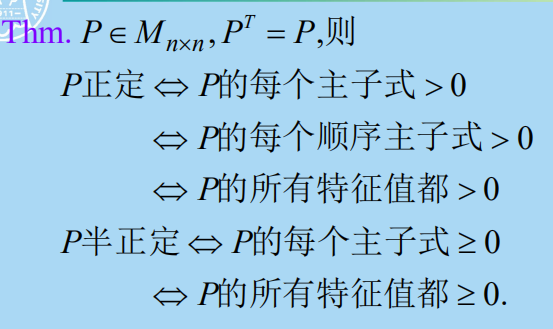


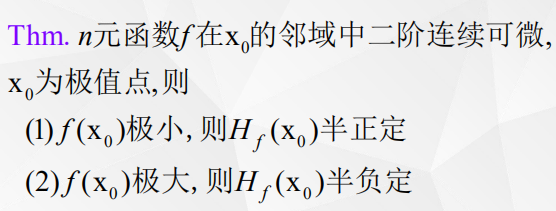
实际上是这句话的扩展，也就是说，你的泰勒多项式必须保留且只能保留到，所有的高阶无穷小都会被放入，所以所有带有的都是高阶无穷小；

极值原理

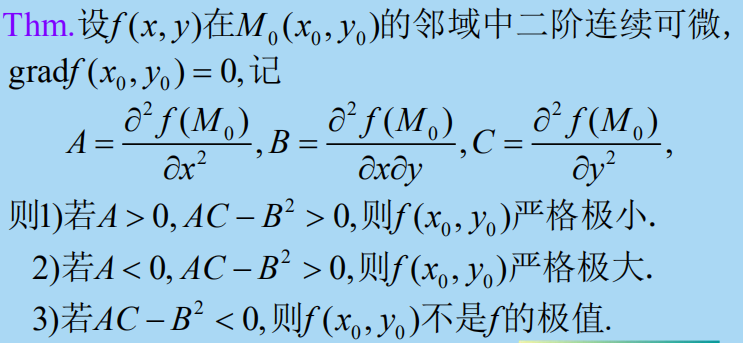
4月14日20:30







实际上，最佳的判据是：



第一个就是每一阶的顺序主子式都为正，那么严格极小；

第二个，奇次阶顺序主子式为负，偶次阶为正，那么严格极大；

（偶阶矩阵取负不影响det；负定矩阵取负为正定）；

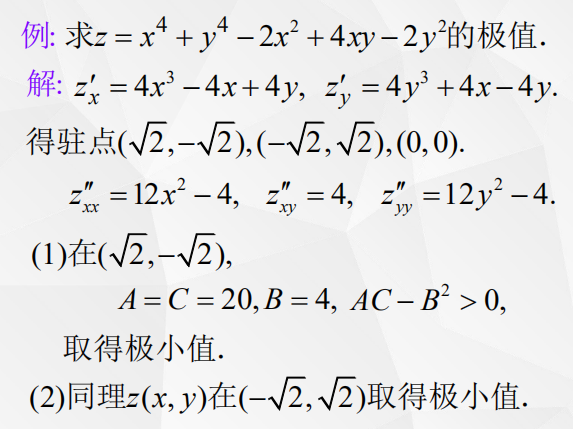
第三个，det为负，必然不定；

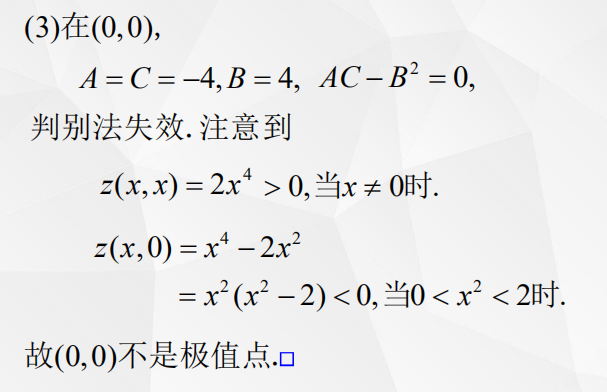
当det=0的时候，不能够根据trace判定半正定性从而判定极性；

因为，极小值推出半正定，但是反之不成立；

故而这种情况下，只有具体问题具体分析；

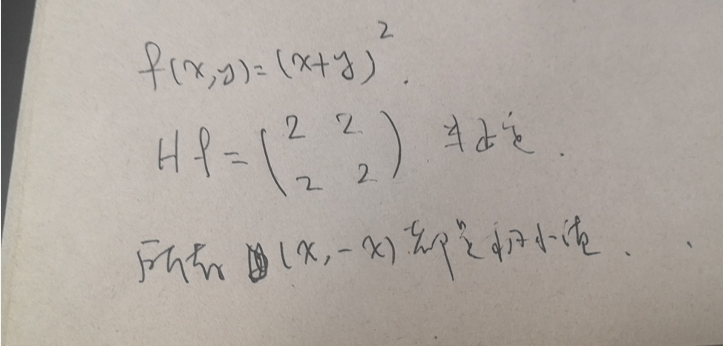
典例：



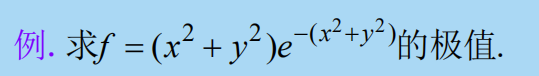


具体分析的方法：

A.退化法：二元函数可以完全退化为一元函数，比如；



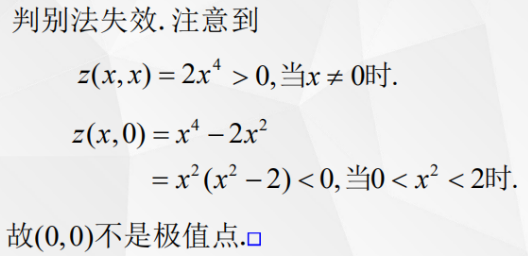
这个二元函数，直接退化为t=x+y的一元函数，对其进行判定即可；

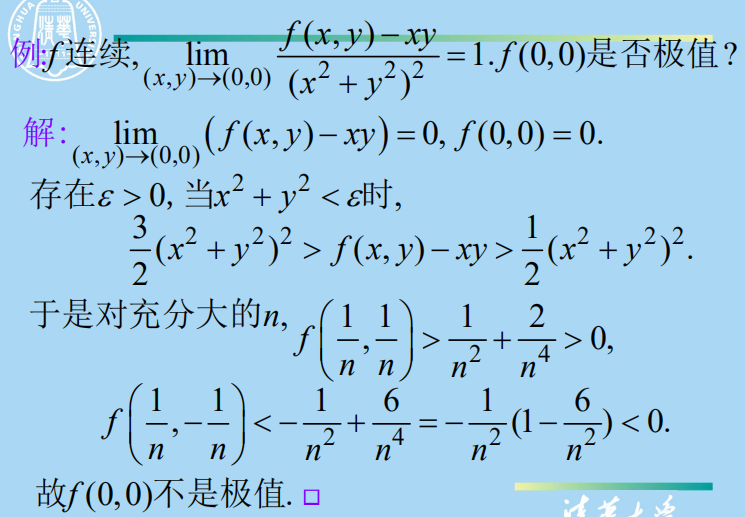


先不换元，求出驻点与黑塞矩阵，如果黑塞矩阵判定不了，再退化为一元，判定一元的极值；

B.特殊路径法：取出det=0的点附近的特殊点来趋近他；如果有的比他大， 有的比他小，那他就不是极值；

常见的特殊路径：f(x,x);f(x,-x);f(x,0);f(0,x)

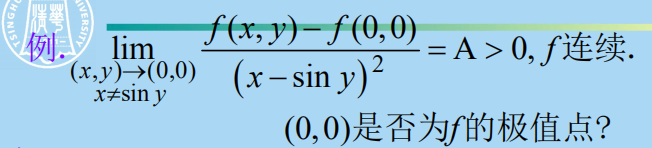




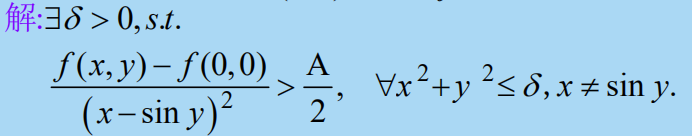
这道题还是特殊路径法，首先通过极限求出f(0,0)必为0；之后圈定范围ε，这种圈定范围的做法很常见，为的就是可以利用，接下来，选择路径法；其实这个特殊路径并不是很明智，完全选择f(x,x);f(x,-x);f(x,0);f(0,x)就可以了；当然这道题里后两个值没有用；

//这里给出的极限有两个作用：第一，给出f(0,0)这一点的函数值；第二，给出在（0，0）附近可以控制住f（x，y）的两个已知函数，由这两个已知函数的特殊路径，确定f（x，y）的特殊路径

对比下两道题：



这两道题的形式都非常相似，而且证明方法也相似；

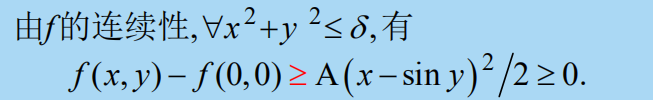


这一步几乎是固定的，通过确定ε来使用极限并且固定要放缩。比如这里是等于A，那就放缩到A/2。上一道题是把1向上放缩到3/2，向下放缩到1/2；

（为什么可以圈定一个极限点的邻域让这个邻域内所有点都满足一个比极限值稍微宽松点的条件呢？因为f的连续性！！！）

（这道题我们并不知道f(0,0)的确切值是多少，所以极限中出现了f(0,0),只需要证明f(x,y)和f(0,0)的差恒大于0就可以了。）

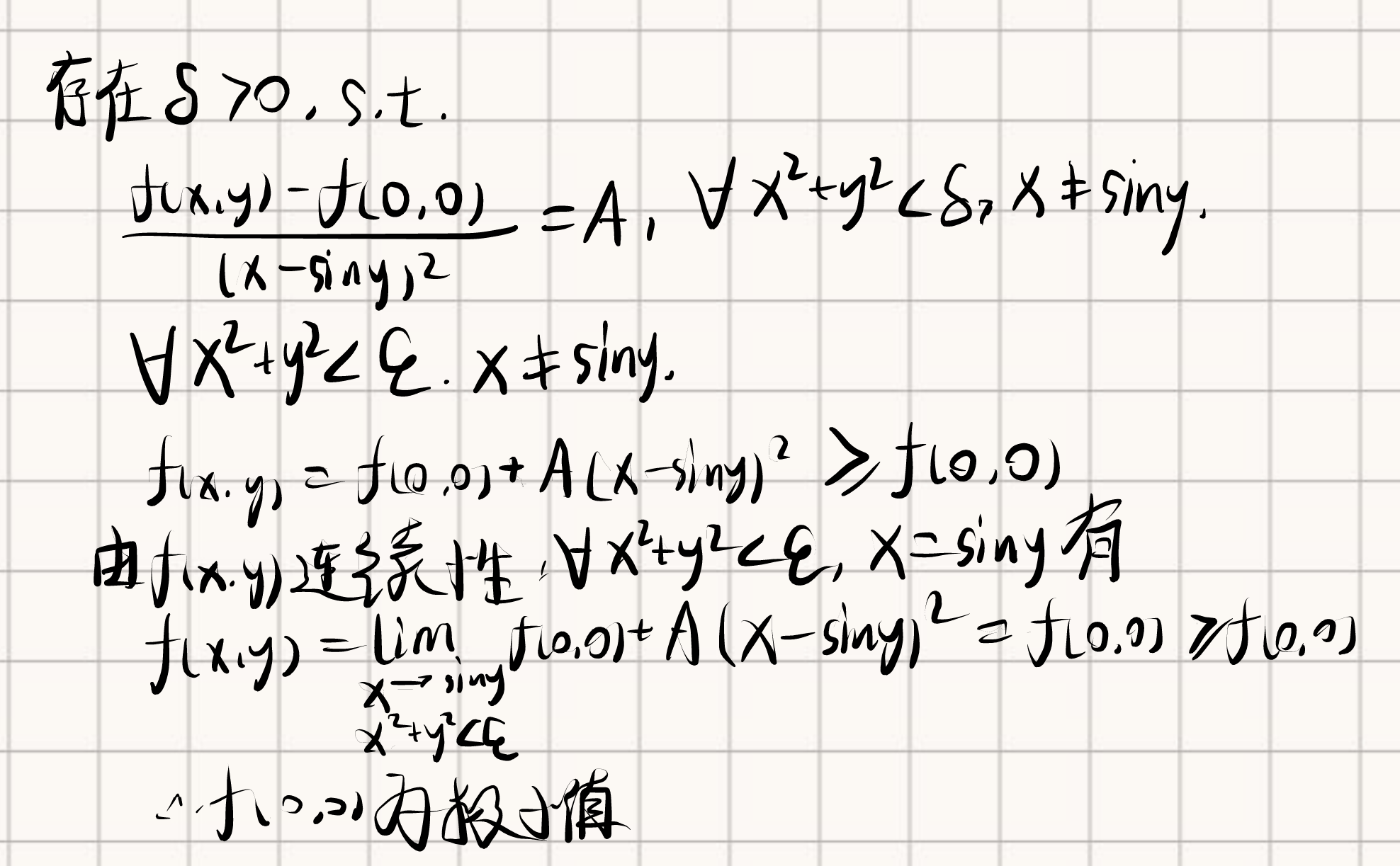
（二重极限隐含的就是“取遍所有路径”）



接下来利用连续性，要么证明极值点；要么特殊路径证明不是极值点；

注意到极值点的定义都是>=或者<=，严格极值点才是>与<；

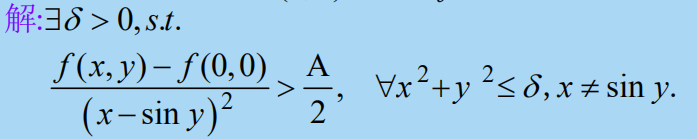
warning：



这里有什么问题？

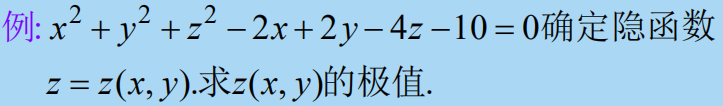
问题就在于，误解了极限的定义；比如我们说1/x的极限是0，但是你能说在某个区间里1/x等于0吗？

所以这个式子是错的，但是

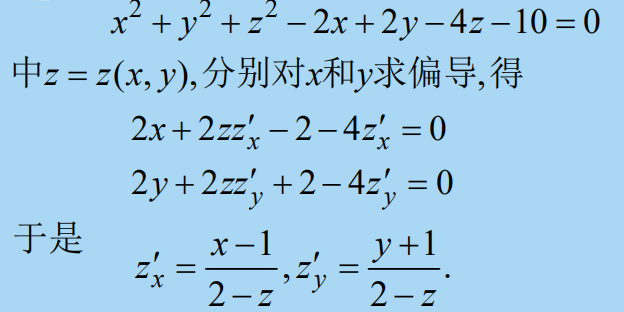


这个一定是对的；所以说你必须要放缩；且可能会同时用到向上与向下放缩；

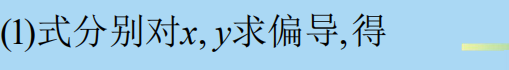
典例：隐函数极值

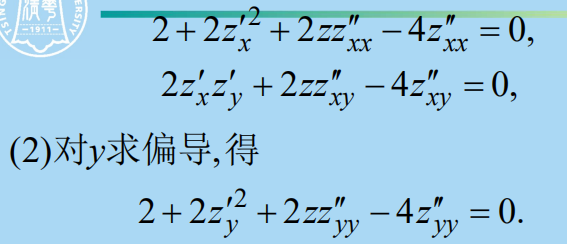


其实还是超平面的想法，驻点和黑塞矩阵都和whatever（这个问题里面是10）无关。然后求出驻点和黑塞；

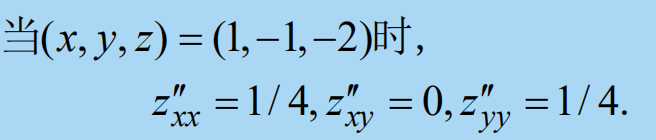


其实这里的做法不理智，理智的做法是直接把俩偏导数为0带进去解方程组；不要整理为分式形式，更快；





接下来求出黑塞；操作都很基础；



但是注意到这里求黑塞的时候，必然要带入x,y,z具体的值，这个就和你的whatever有关了。

只能说，whatever影响具体点上的偏导数取值，但是不影响偏导数的表达式。

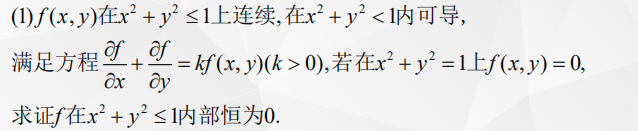
极值原理的证明题

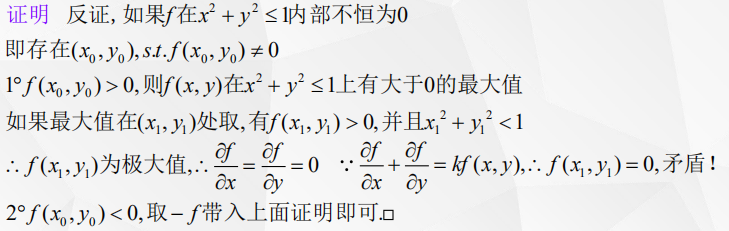
A.题目特征：给定连续函数在某一闭集里Hessian矩阵或者Jacobi矩阵的性质，再给定边界恒为常值，证明区间内部的性质；

B.这类题的套路都很明显：反证法+极值原理

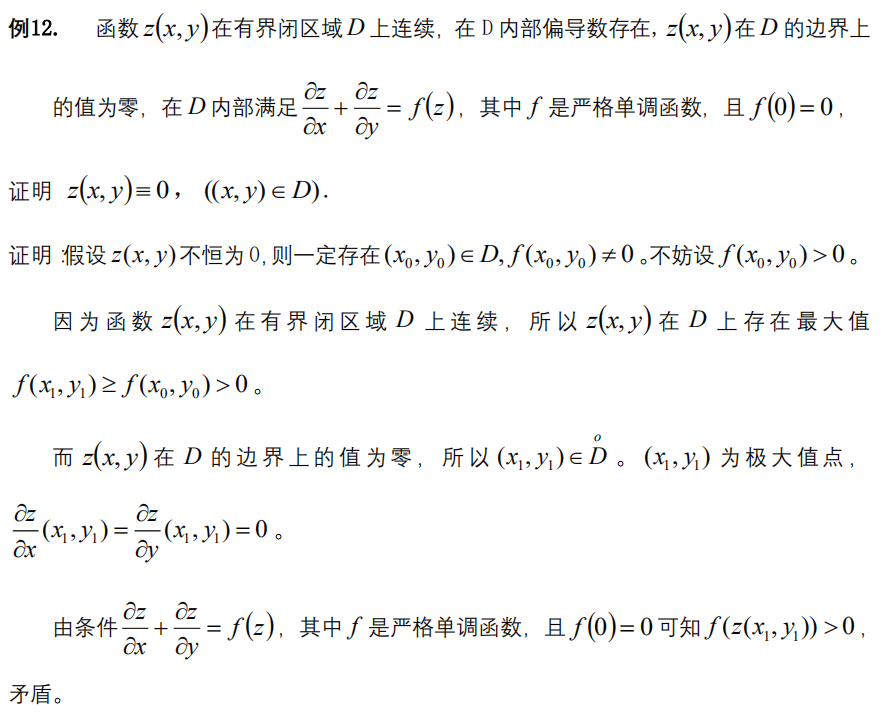
具体而言，假设存在反例，那么一定可以判定出最值点在区间内部。（因为边界为常值）最值点在区间内部就必然满足Jacobi为0，Hessian也有相应的（半）正定（半）负定性；结合已知条件利用det、trace、甚至梯度就能导出矛盾；

第一题：

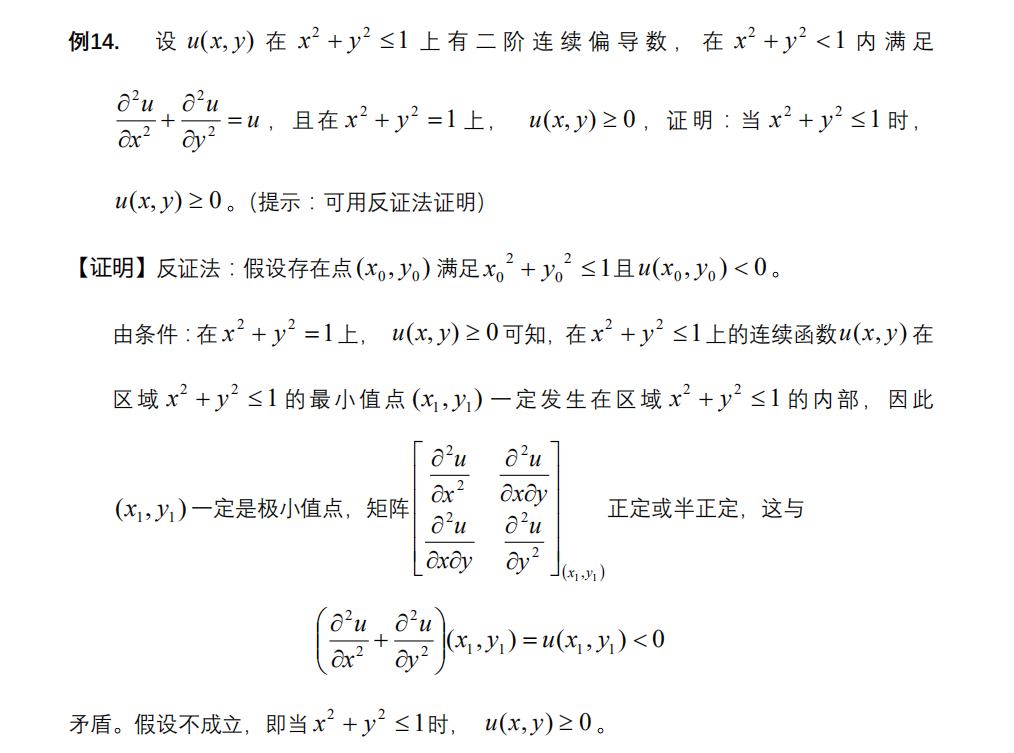




第二题：

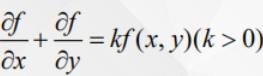
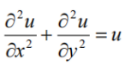
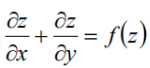


第三题：



你会发现，这三题几乎一模一样；

当然，还是要强调下，需要读懂条件：

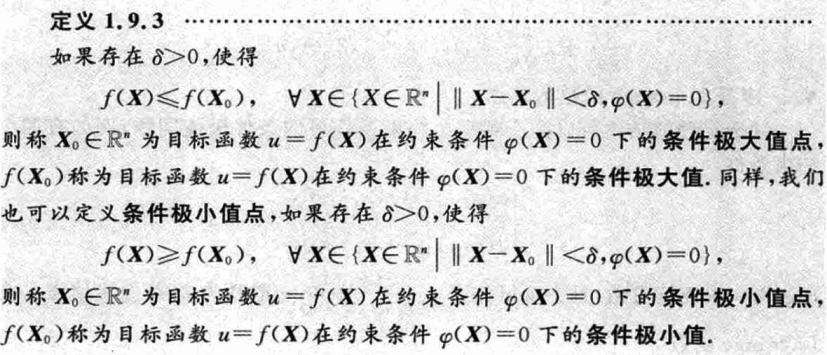
这三个条件都是描述某一点的偏导数和某一点的函数值的关系；

条件极值

一、问题定义：

我们虽然把这个问题叫做条件极值，但是条件极值，根本就不是不加条件下的原问题的极值，也不是Lagrange函数的极值。

这里先理清楚几个概念，最值，极值，条件最值，条件极值；



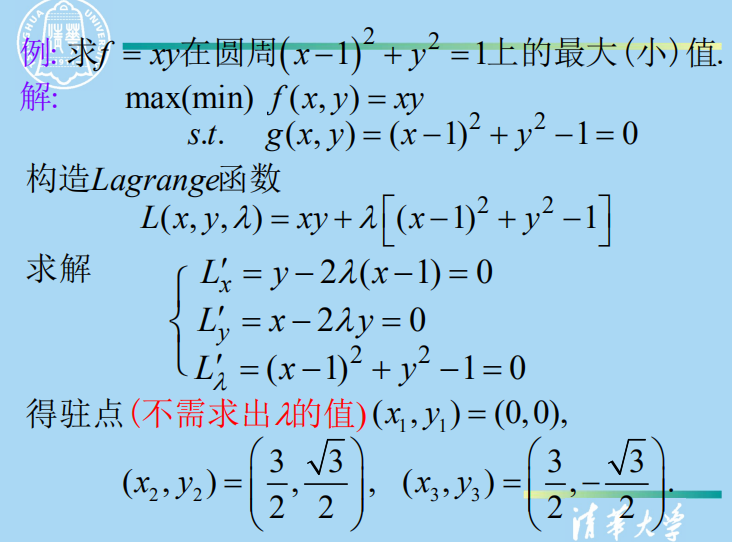
条件极值求出的Xo（其实不用求出λ），仅仅是拉格朗日函数的一系列驻点。

条件极小值或者条件极大值中的某一个（或者两个，但至少有一个，常常另一个不存在）一定在这些驻点之中。然后原问题在这一限制条件下的最大值或最小值中的某一个或者两个一定也在这些驻点之中，但完全不知道是哪个驻点，也不知道究竟是最大值还是最小值。

有一个驻点，就有条件极值，就有条件最值；但是有多个驻点，不一定有两种条件最值；因此，完全可以归纳为：

得到驻点后，你只是得到了疑似的最大或最小值点，需要自行判定这些点，是最大，还是最小，还是啥都不是。但是最大最小值中的一个一定在这些驻点中的某一个取得，另一个最值甚至不一定能取得，或者说往往不存在。

比如经典的例子：



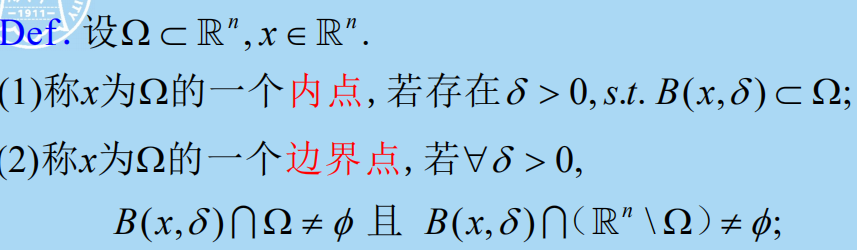
你得到了三个点，这三个点现在都只是有嫌疑，还完全不知道究竟是不是这一条件下的最值。需要人为判定，那么怎么判定呢？

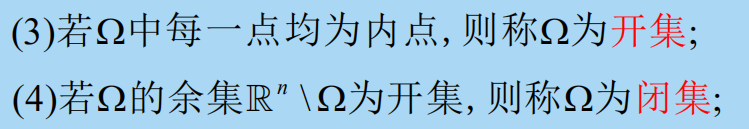
常见的判定方法有几种：

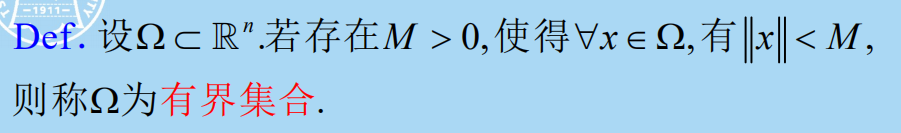
二、判定方法：

A.有界闭集法：

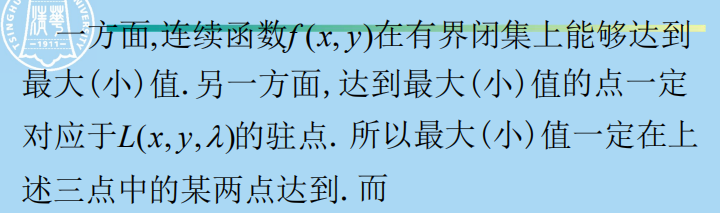
首先，定义，什么是有界闭集？





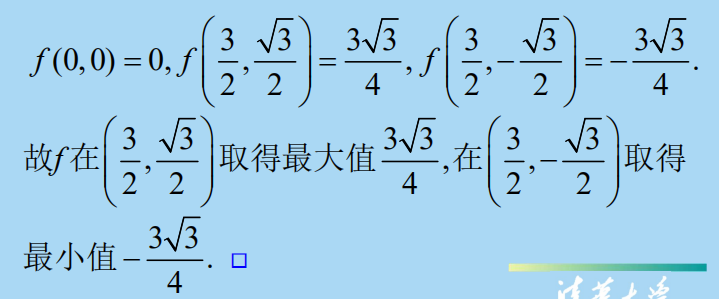


注意到上面那个例子：

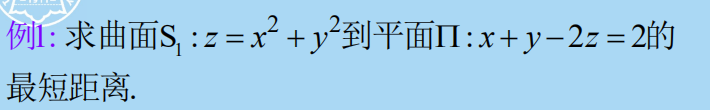


圆周的余集为空集，空集既是开集也是闭集，开集的补集是闭集，那么圆周是闭集。同时圆周上的点到原点的距离有界，那就是有界闭集。

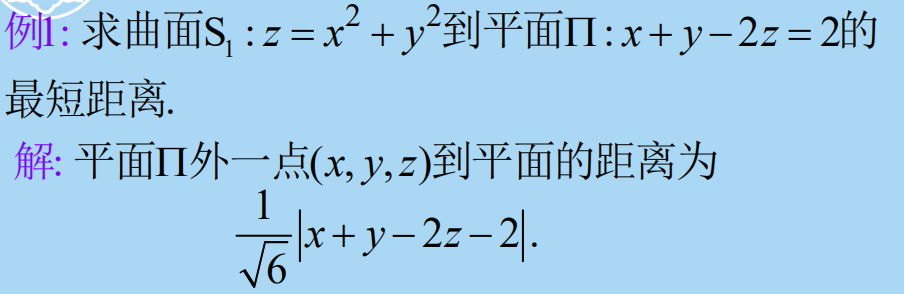
利用有界闭集的性质得出一定存在最大最小值点，那么这俩最值一定是这些驻点中的某两个，一个一个带进去试一试就可以了。



B.直观说明法：



首先，这道题默认你会到平面的距离公式；



到直线的距离公式呢？（我已经忘了，呜呜呜）

其实类比下很好出来的；

anyway吧，注意到这个题，我们需要分清楚，谁是约束条件，谁是目标函数，显然约束条件是，注意到，这个不是一个有界闭集；他是无界闭集；那么目标函数在无界闭集上无法得知有无最大最小值；另一方面，拉格朗日函数有驻点，那么目标函数在约束条件下必然有最值。

接下来，对目标进行优化；

对目标的优化往往非常关键，比如距离函数的常见优化就是处理平方项，而的优化是取对数；



剩下的计算问题不大，只是需要注重说明究竟是最大还是最小；

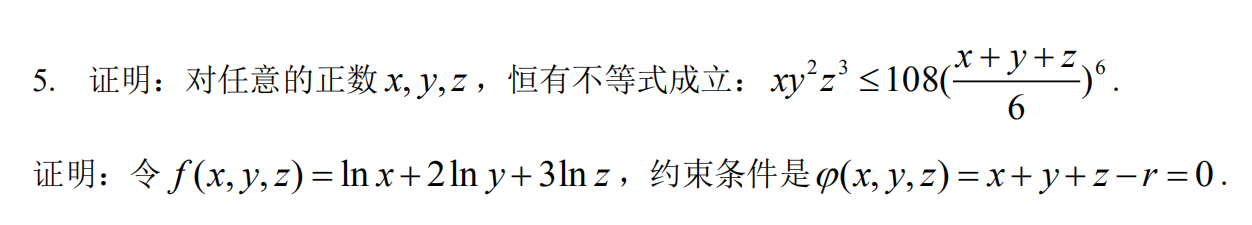
//这里的计算技巧：用1（或者2）表示lambda，用3也可以表示lambda，两式联立可以直接得到x=0或x=1/4.

注意到z=0，x=-y时，目标函数可以取得无穷大的值，故而目标函数在该约束条件下不存在最大值，那么拉格朗日函数的唯一驻点必然是最小值；

二、使用技巧：

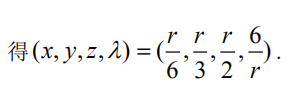
A.目标优化:

前文叙述过，条件极值的优化是个非常重要的技巧：

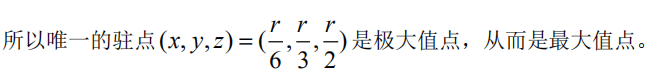


通过取对数优化了目标函数，而对左边，直接令x+y+z为常数值，也进行了一定的优化。

之后得到了唯一的驻点；

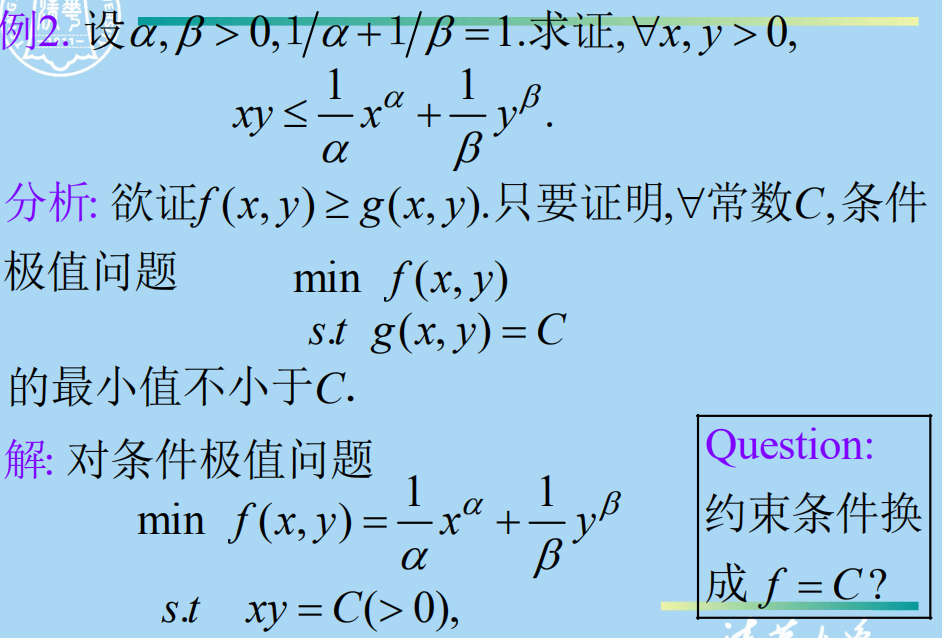


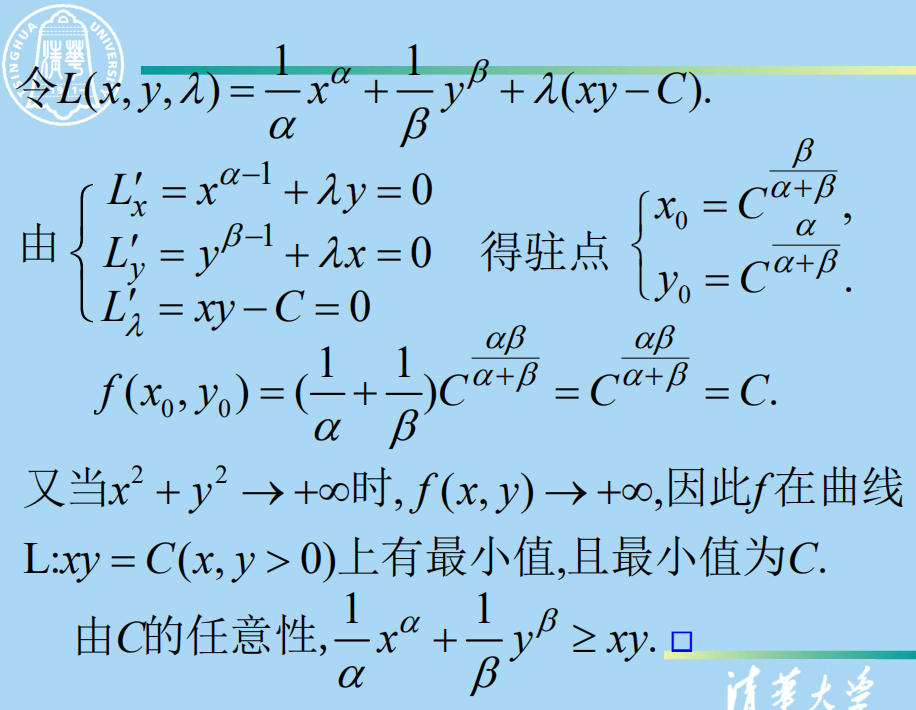
我们可以用直观说明法说明目标函数在该约束下没有最小值，而得到了唯一驻点；那么：



tips：此题当然不能用有界闭集法，因为你左侧显然不有界；

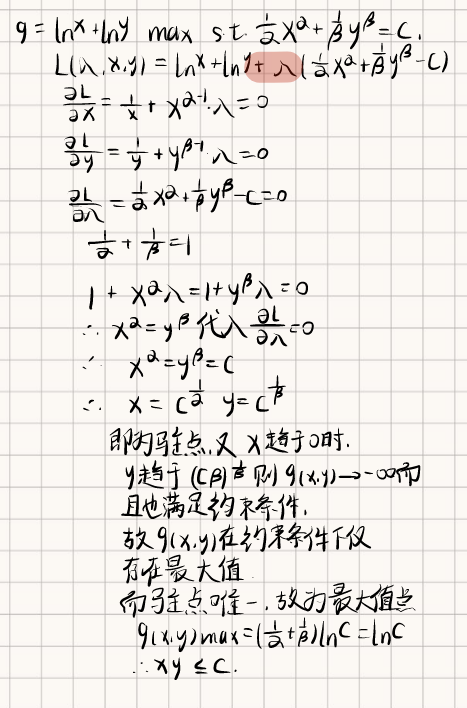
同样的，这道原题：





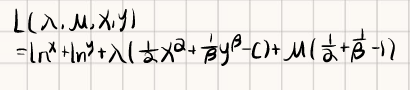
原答案是将不等式左边视为了约束条件，操作了右边，并结合了直观说明法得出了唯一驻点为条件极小值点，也是条件最小值点。

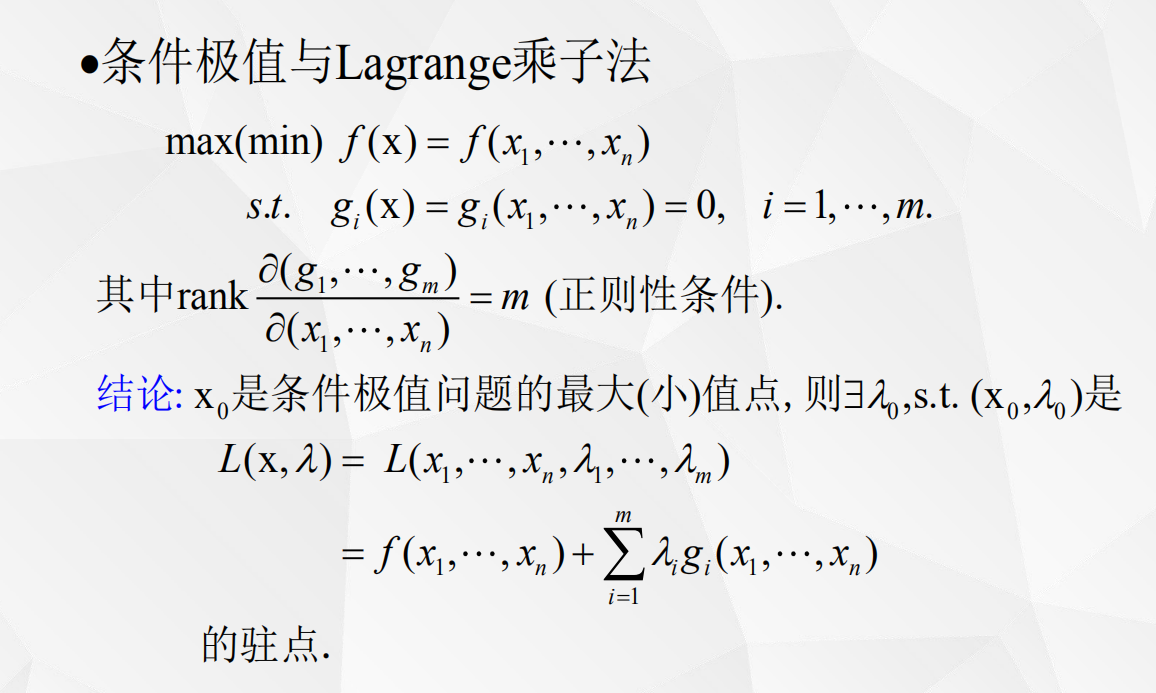
其实进行目标优化之后：也可以实现；



B.多约束极值：

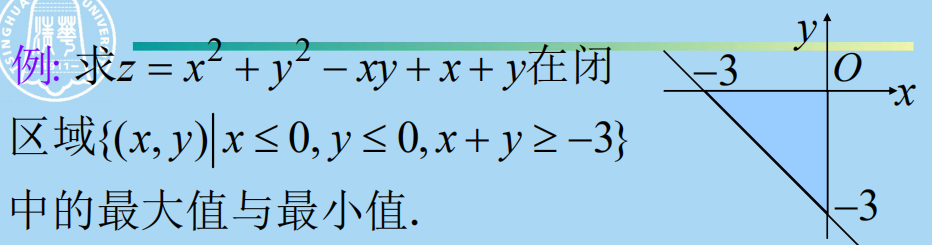
实际上，上述的例题可以视为一个多约束问题：





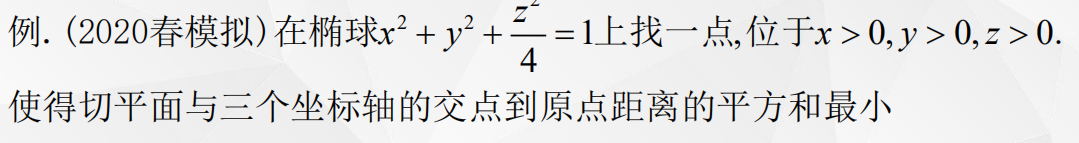
如法炮制即可；

C.闭区间最值问题：

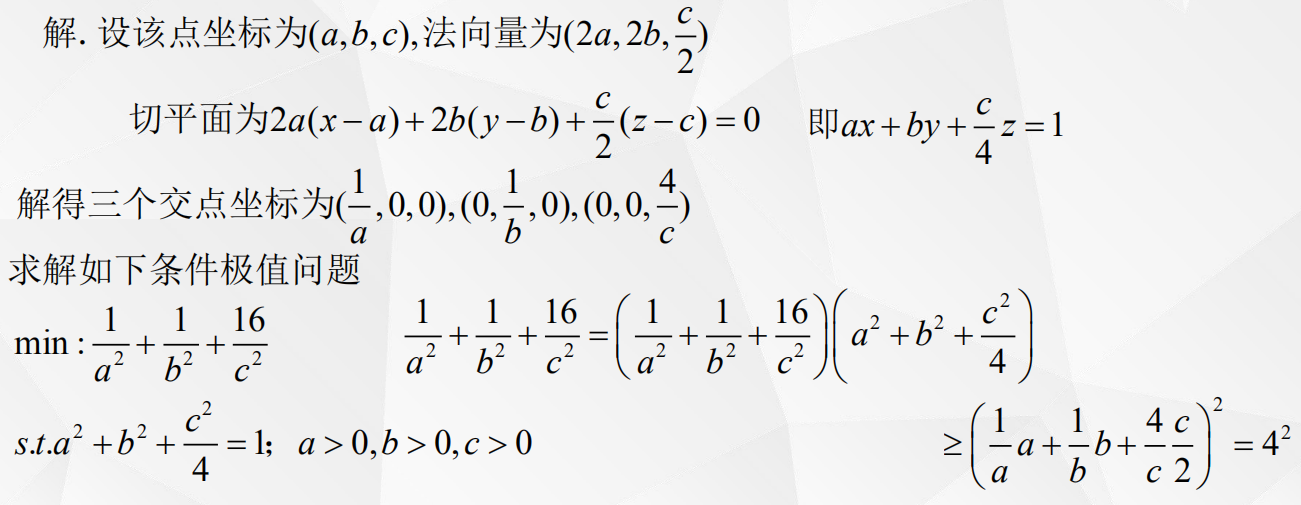


这种题无非就是先求出Z在区域内的所有驻点（不要判定黑塞，因为极值不一定是最值），接下来在边界上求出边界上的最值；（由于太过简单，直接消元就可以了），然后比较边界和几个极值，其实完全不用判定黑塞；

例题：

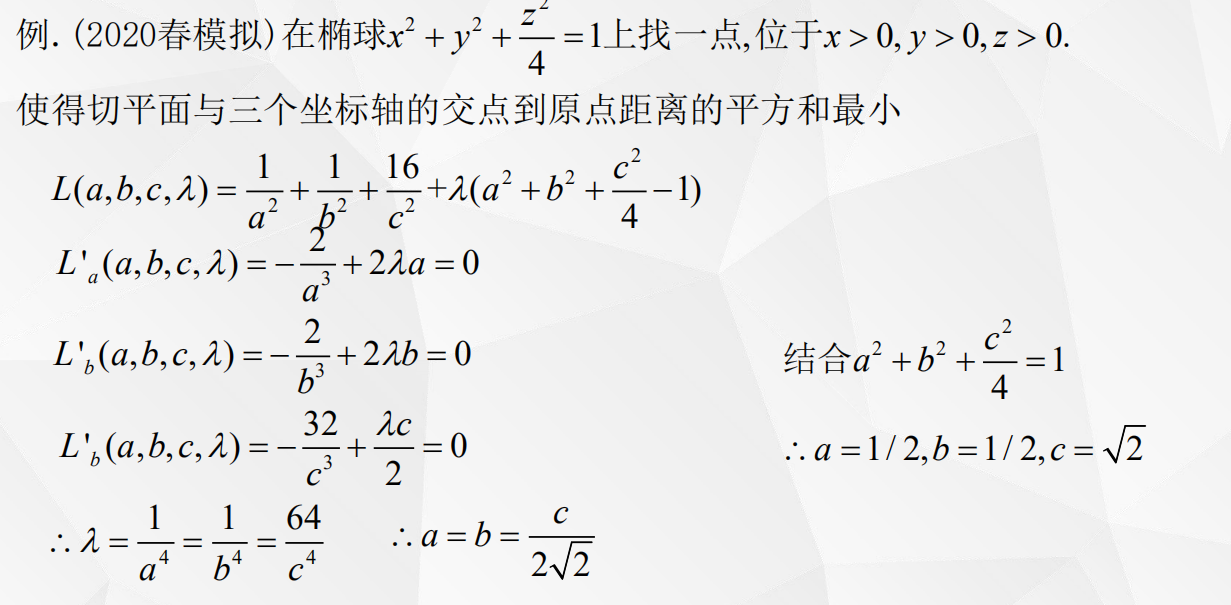


老老实实求出交点；

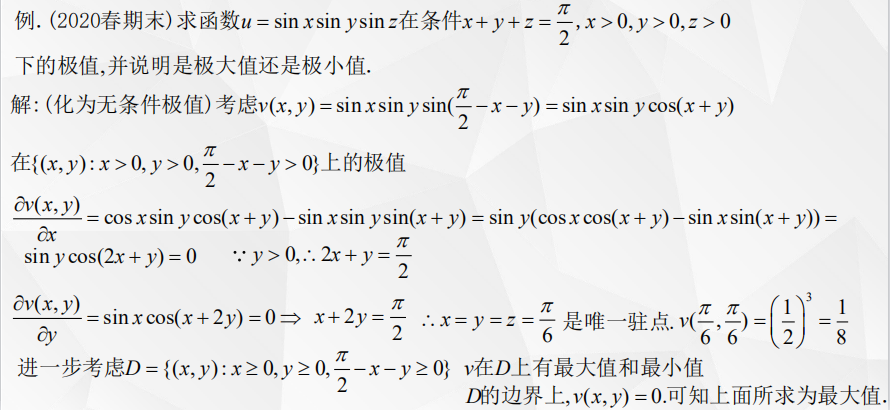


注意到，求出切平面的那一步带入了椭球方程；

这里使用了柯西不等式，当然很巧妙，不过考试不一定（我觉得是一定不）希望你这么做；



使用拉格朗日乘子法得出唯一驻点；然后由于目标函数一定可以无限大，所以这一定是最小值；



最基本的方法，把条件极值转化为区域最值；其实我觉得这里已经足以说明问题了，因为最大值必然在内部，内部只有一个驻点，那驻点必然为极大值；但是还是可以验证下黑塞说明问题；