第七章.多元函数微分学

表面看来，二元函数只是多了一个自变量，有关的微积分概念应该是与一元情形相平行的，似乎是不需要多少新的概念，实际上，二元的情形从几何的意义来看，就是从一维进入了二维，而我们知道二维的几何现象要比一维的几何现象丰富得多。因此有关二元函数的微积分概念注定要比相应的一元情形复杂得多。也就是说出现了很多新的问题与新的现象，需要应用新的概念来给予刻划。

二元函数及其定义域。

对于二元函数的自变量，我们可以用两种观点来看。一就是两个相互独立的自变量，分别具有其变化范围，要确定一个函数值，就必须同时取定两个独立的自变量的取值，为了获得比较直观的看法，可以首先假设一个变量取某个定值，然后考虑剩下的一元函数，这样对于不同的定值，就有相应的一元函数，再反过来，考虑另一个变量，类似地得到对于整个函数的一个大概了解。另外一种观点，就是把两个自变量看成是一个平面上点的两个坐标分量，这样自变量的变化就表现为这个点在平面的位置的变化。这时我们就有可能在平面上给出函数的定义域的描述，一般情况下，就是一定的平面区域。这样函数就可以想象为投影为定义域区域的三维曲面。

第二种观点具有很强的直观性，因此我们对二元函数的研究，主要就采用这种看法。

给出一个二元函数的解析式以后，确定它的定义域是分析函数的关键第一步。而初学者往往正好是忽略这一步，认为过于简单而无关紧要，实际上给出某些函数的定义域的平面图形还是非常麻烦的一件事，并且我们强烈要求初学者，不管函数多么简单，都一定要在进行任何的处理之前，首先分析它的定义域，画出它所表示的平面区域的草图。

有关定义域所表示的平面区域，具有两个相当重要的概念：开闭性和连通性。而基本概念仍然是邻域的概念。

平面区域的开闭性非常复杂，理解它的基础就在于邻域的观念。严格的定义超出了本课程的要求，我们只要求掌握一些常见的函数的定义域的开闭性质，初学者特别应该加强训练，多作相关练习，养成在处理二元函数时，总是先动手画出函数定义域的平面图形的习惯。

连通性的概念也非常重要的，因为我们有可能遇到一些定义域由不连通的区域所组成的情况，这就增加了分析问题的复杂性，也是常常容易发生错误的地方。

二元函数的极限。

对二元函数定义极限以及连续性，关键是定义变量从一个值趋向另一个值的含义，前面我们已经提到对于二元函数的两种看法，这两种看法导致了不同的两种极限。

首先我们看和一元函数类似的第二种看法，即把两个自变量看成平面上一个点的两个坐标分量，这样自变量从一个点趋向于另一个点的行为可以理解为到另一个点的距离趋向于0，而平面上点之间的距离是具有现成定义的，即对于平面上的两点（a，b）和（c，d），它们之间的距离定义为

$$l=\sqrt{(c-a)^2+(d-b)^2}$$

运用这个定义类比于一元函数的坐标差的绝对值的距离的定义，就可以得到二元函数的极限的相应定义。

这里我们强烈建议同学们根据这个提示，自己写出二元函数的极限的严格定义，只有自己能够写出来，才能避免阅读时看得懂，而当要求自己叙述时，却又理不清头绪的问题，特别是这样能够加强我们对于极限这个微积分基本概念的理解。

进一步，对于上面的观点，我们还可以完全基于邻域的概念来理解，因为在邻域的概念里，已经包含了距离的使用，因此，直接使用邻域的概念，就可以非常简洁地得到与一元函数统一的定义，实际上，也就是对于任意的多元函数的极限的统一定义：

设函数f（X）在某点P的某个去心邻域有定义，如果对于任意的$e>0$，都存在相应的$d>0$，使得当点X属于这个邻域时，总有

$$|f(X)-A|<e$$

成立，我们就称f（X）在X趋向于P点时，以A为极限。

这里点X或者点P的坐标分量有多少个是没有具体限制的，这就实际上是给出了对于任意的多元函数的极限的定义。并且可以称为是多元函数的多重极限。

现在，我们再讨论如果使用对于二元函数的第一种看法，会给极限定义代来什么样的变化。

如果我们分别考虑二元函数的两个变量，那么对于函数z=f（x，y），自变量（x，y）的取值趋向于某个定值(x_0,y_0)，就可以看成是首先把一个变量，例如y看成参数，只是考虑变量x的趋向于x_0点的行为，这时，我们可以完全应用一元函数的极限的定义，得到

$$\lim_{x\to x_0} f(x,y)=g(y)$$

的定义，然后，再考虑包含变量y的g（y）的相应极限行为，同样应用一元函数的定义，得到

$$\lim_{y\to y_0} g(y)=A$$

的定义，这样，我们实际上是定义了这样一个极限过程：

$$\lim_{y\to y_0}\lim_{x\to x_0} f(x,y)=A$$

。

这被称为二次极限，一般地，这样定义的多元函数的极限被称为累次极限。

初学者千万需要注意的是，累次极限与多重极限具有细微的区别，它们并不是完全的等价，而是具有下面的定理所表明的不等价关系：

如果函数f（x，y）在点(x_0,y_0)的二重极限为

$$\lim_{\substack{x\to x_0\\ y\to y_0}} f(x,y)=A$$

（无论是有限还是无限），

同时，如果对于任一邻近y_0的y，存在关于x的有限极限

$$\lim_{x\to x_0} f(x,y)=g(y)$$

，

那么二次极限$\lim_{y\to y_0}\lim_{x\to x_0} f(x,y)=\lim_{y\to y_0} g(y)=A$存在，并且是等于二重极限A的。

从什么的定理，可以看到多重极限与累次极限的关系是比较复杂的，对于具体的问题需要作具体的分析，而不能武断地用一个代替另一个。特别是由于计算累次极限实际上就是计算一元函数极限的过程，比较计算多重极限要简单，使得初学者特别容易在这方面犯错误。

二元函数的连续性及其性质。

定义了二元函数的极限以后，对于连续性就容易理解了。

类似于极限定义当中，我们分别应用了点之间的距离和邻域的观点，相应的对于连续性，同样可以使用这两种观点进行叙述。

我们在讨论一元函数的极限与连续性时，已经知道极限与连续性的重大区别在于极限不要求函数在极限点有定义，而连续则一定要求函数在该点有定义，并且函数在该点的极限必须就等于函数在该点的函数值。

因此相应的，我们在定义二元函数的连续性时，同样要强调函数在极限点处有定义，然后就是

$$\lim_{X \to P} f(X) = f(P),$$

就定义了函数在P点的连续性。

上面定义的是函数在某点的连续性，相应的，函数在某个区域的的连续性，就是函数在这个区域的每个点的连续性，而对于连续函数经过四则运算与复合而得到的函数，在最终的定义域内，同样是连续性，因此从本质上来看，四则运算和复合都不影响函数的连续性。这些都是和一元函数平行的结果。

但对于函数的间断情形，二元函数就要比一元函数要复杂，不过产生间断的原因则是同样的，即或者是由于函数在某点不存在定义，或者是虽然存在定义，但在该点的函数值不等于函数在该点的极限值。

还有一个与一元函数完全平行的连续性性质，就是连续函数在闭区域上的最值性质与介值性质：

（1）如果函数f（X）在有界区域D上连续，则必定存在点 $P_1, P_2 \in D$ ，使得对于任意点 $P \in D$ ，都有

$$f(P_1) \le f(P) \le f(P_2)$$

成立，我们相应地称P_1为函数在区域D的最小值点，而f（P_1）为函数在区域D上的最小值，相应的就是最大值点和最大值。这个定理被称为最值定理。

（2）如果函数f（X）在有界连通闭区域D上连续，a，b为函数在区域D上的任意两点的函数值，则对于区间[a，b]上的任意实数c，必定存在区域D上的一点Q，使得

f（Q）=c。

这个定理被称为介值定理。

注意这两个定理的条件的重要差异，即介值定理要求是在一个有界连通闭区域，而最值定理只是要求在有界闭区域，也就是说，介值定理要求函数定义区域的连通性。

偏导数。梯度。

对于多元函数，同样需要考虑因变量关于因变量的变化率，也就是类似于单变量函数的导数的概念，但是由于多元函数的自变量是由多于一个的几个变量组成，因此情况变得比较复杂，这样我们只有首先尝试在讨论极限时已经运用的方法，就是除了一个变量以外，其他变量都看成是参数，不考虑它发生变化的情况，也就是取固定值的情况，这样，就得到了偏导数的概念。

偏导数的定义是完全平行于单变量函数的导数的定义的，在这里，除了要取偏导数的那个自变量以外，其他的自变量都不用管，只当是一个常数就可以了，这样，就可以对函数的每一个自变量单独地取偏导数。对于具有相互独立的n个自变量的函数，就可以定义相互独立的n个各自的偏导数。而求导法完全就是按照单变量函数的求导法，再强调一句，因为在求偏导数时，其他的自变量都可以看成是常数

不过要注意，不要把单变量函数的导数的符号 $\frac{d}{dx}$ 和对于x的偏导数的符号 $\frac{\partial}{\partial x}$ 搞混淆了，毕竟起码它们的定义背景不同。而且它们还有两个重要的差别：

（1） 对于多元函数来说，如果函数在某点的对于各个自变量的偏导数都存在，但仍然有可能函数在这点是不连续的。

（2） 偏导数的符号 $\frac{\partial}{\partial x}$ 是一个整体，它暗示了还有其他的变量被看成是常数，而单变量函数的导数的符号 $\frac{d}{dx}$ ，我们知道根据微分的含义，可以把它看成是一个商式，也就是说，满足平常

的乘除法运算法则，而$\frac{\partial}{\partial x}$则不允许这么理解，因为对于多元函数的偏导数，还不存在类似于单变量函数的微分那样的概念。

按照把其他自变量看成常数的理解，就可以很容易地理解偏导数的几何意义，例如对于一个二元函数f（x，y），把一个自变量y取为常数，就意味着用垂直于Y轴的平面去横截函数f（x，y）所表示的曲面，然后可以得到沿着横截面的曲线在某点的切线，切线的斜率就是函数在该点的对于x的偏导数的几何意义。

正是由于偏导数假定了其他的自变量为常数，因此不能令人满意地表达函数随着所有自变量变化而产生的变化率，如果我们希望定义这样一个能令人满意地表达函数随着所有自变量变化而产生的变化率的新的量，则根据多元函数的几何图象，我们知道函数随着所有自变量变化而产生变化时，是可以有不同的方向的，因此可以想象这个新的概念还必须是一个向量，基于这些考虑，我们给出梯度的概念：

设二元函数f（x，y）在点（x，y）处的偏导数都存在，则函数在该点的梯度为

$$gradf(x,y)=\nabla f(x,y)=\frac{\partial f(x,y)}{\partial x}\vec{i}+\frac{\partial f(x,y)}{\partial y}\vec{j}$$，

其中$\vec{i}$和$\vec{j}$分别是在X轴和Y轴方向上的单位向量。

可以看到这个定义是一个线性组合。因此可以很容易地想象，这个运算具有线性性质，也就是

$$\nabla(f+g)=\nabla f+\nabla g$$；

$$\nabla cf=c\nabla f$$；

$$\nabla(fg)=f\nabla g+g\nabla f$$。

全微分及其几何意义。

类似于单变量函数的微分的概念，对于多变量函数就是全微分的概念。

我们知道单变量函数的微分就是变量增量的线性近似，对于多元函数，首先必须定义相应的自变量增量的概念，为了反映所有自变量的变化所导致的因变量的变化，我们定义全增量的概念，就是每一个自变量产生一定的增量，相应地给出因变量的增量，这个增量是由所有自变量的变化而产生的，因此称为全增量，即对于函数$z=f(x,y)$，全增量就是

$$\Delta z=f(x+\Delta x,y+\Delta y)-f(x,y)$$。

定义了全增量，那么顺理成章地，类似于单变量函数的微分的定义，我们把全增量依赖于每一个自变量增量的关系线性化，并且把全增量的线性主部就称为函数的全微分：

$$\Delta z=f(x+\Delta x,y+\Delta y)-f(x,y)$$

$$=A\Delta x+B\Delta y+o(\sqrt{(\Delta x)^2+(\Delta y)^2})$$。

其中定义全微分为

$$dz=A\Delta x+B\Delta y$$。

那么我们就可以把函数在某点处可微定义为函数在该点的某个邻域存在定义，而基于该点的全增量可以表示为上面的形式，也就是说明可以定义全微分。

根据可微的定义，可以直接得到下面的定理：

如果二元函数f（x，y）在（a，b）点处可微，那么

（1） 函数在（a，b）处连续；

（2） 函数在该点处的各个偏导数都存在，并且函数在该点处的全微分满足

$$df(a,b)=\frac{\partial f(a,b)}{\partial x}\Delta x+\frac{\partial f(a,b)}{\partial y}\Delta y,$$

或者写成

$$df(a,b)=\frac{\partial f(a,b)}{\partial x}dx+\frac{\partial f(a,b)}{\partial y}dy,$$

如果直接利用梯度的向量表达式的优越性，也可以直接写成

$$df(\vec{x})=\nabla f(\vec{x})\cdot d\vec{x}\text{。}$$

在这个表达形式里，可以看到梯度于导数的类似性。

进一步，我们有如下的有关偏导数与可微的关系的定理：

如果对于函数在某点的某个邻域内，所有的偏导数都存在并且连续，那么函数在该点必定可微。

注意这个定理里关于可微的充分条件。

我们知道二元函数能够用来描述一个三维空间中的曲面，而一个二元函数在某点可微，则刻划了这个二元函数所表示的曲面在该点的一个重要性质，即在该点的切平面的唯一存在性，定理如下：

如果函数 $z=f(x,y)$ 在点（a，b）处可微，则相应的曲面在点（a，b，c）存在唯一的一个切平面（$c=f(a,b)$），切平面方程为

$$z-c=f_x(a,b)(x-a)+f_y(a,b)(y-b)\text{。}$$

可以看到这个定理实际上表明了可微的几何意义，即和单变量函数的导数的意义一样，只是对于多变量函数来说，偏导数并不具有和单变量函数的导数一样的地位，而只有全微分才具有和单变量函数的可微（对于单变量函数来说，就是可导）相同的地位。

方向导数。

我们已经知道描述多元函数的在某点处的一般变化率的是梯度，而梯度是一个向量，因此它在某个确定的点处是具有确定的方向的。而在实际应用当中，我们不只是需要知道函数在梯度方向的变化率，也还要求知道其他特定方向的变化率，这种根据特定方向而计算出来的变化率，称为方向导数。

具体到二元函数，就可以想象是在曲面上，对于自变量所张成的空间内的一个方向向量 $\vec{l}$，在某点处沿着这个方向求导数，即是这点处在这个方向上的方向导数。记为 $\dfrac{\partial f(x,y)}{\partial \vec{l}}$。

考虑到梯度的定义以及向量的内积的几何意义，就是一个向量在另一个向量方向的投影，因此很自然可以得到函数在某点的方向导数就是在，该点的梯度在所给的方向上的投影。在就是下面的定理：

如果函数 $z=f(x,y)$ 在点（x，y）处可微，则函数在该点沿着任意方向 $\vec{l}$ 的方向导数都存在，并且等于

$$\frac{\partial z}{\partial \vec{l}} = \nabla f(x,y) \cdot \vec{l}$$ 。

从这个具有明确几何意义的定理，可以得到梯度的两个几何性质：

（1） 函数在某点的梯度向量与过该点的函数的等值线正交；

（2） 函数在某点的梯度向量的方向正是函数值在该点处变化最快的方向。

多元复合函数的求导以及微分形式不变性。

通过四则运算而构造的多元函数的偏导数，具有与单变量函数一样的四则运算法则，而对于复合过程，则比单变量函数要复杂，但仍然是运用类似的链导法，根据函数复合的方式的不同，链导法也分成几种情况。

（一）全导数。

所谓全导数本质上还是单变量函数的导数，即对于一个单变量函数，中间自变量却出现了两个，这样函数对于最终自变量的导数就是全导数，计算公式为

设我们有函数 $z = f(x,y)$ ， $x = x(t)$ ， $y = y(t)$ ，那么

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

写成向量形式就是

$$\frac{dz}{dt} = \nabla f(\vec{x}) \cdot \vec{x}'(t)$$ ，

其中 $\vec{x}$ 是具有分量形式（x，y）的向量。

条件是各个函数在相应的点处都是可微的。

这个计算公式还可以很自然地加以推广到任意有限多的中间变量。

（二）偏导数。

如果最终变量和中间变量都是两个，则有下面的计算方法：

设有函数 $z = f(x,y)$ ， $x = x(s,t)$ ， $y = y(s,t)$ ，而某点（s，t）处函数 $x = x(s,t)$ 和函数 $y = y(s,t)$ 的所有偏导数都存在，而函数 $z = f(x,y)$ 在相应的点处可微，那么

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}$$ ；

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$$ 。

与微分对于单变量函数的复合方程来说，具有微分形式不变性一样，全微分对于函数的复合过程也具有微分形式不变性。通过进行链导法运算就可以得到：

设有函数 $z = f(x,y)$ ， $x = x(s,t)$ ， $y = y(s,t)$ ，则复合函数 $z = f[x(s,t), y(s,t)]$ 的全微分dz为

$$dz = \frac{\partial z}{\partial s}ds + \frac{\partial z}{\partial t}dt$$

$$=(\frac{\partial z}{\partial x}\frac{\partial x}{\partial s}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial s})ds+(\frac{\partial z}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial t})dt$$

$$=\frac{\partial z}{\partial x}(\frac{\partial x}{\partial s}ds+\frac{\partial x}{\partial t}dt)+\frac{\partial z}{\partial y}(\frac{\partial y}{\partial s}ds+\frac{\partial y}{\partial t}dt)$$

$$=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy$$ 。

可以看到无论是对于中间变量还是对于自变量，全微分的形式都是一样的，因此这就称为微分形式的不变性。

隐函数的求导。

利用对于复合函数的链导法，就可以对隐函数进行微分，因为通过对于隐函数，我们总是可以随时通过引入中间变量，而进行微分。

所谓隐函数形式实际上就是可以看成是方程的形式，而表示函数的方程一般有使用单个方程的形式，或者说使用方程组的形式，其中讨论单个方程是基础。

（一）对于单个方程，首先是要确定哪个变量作为自变量，而哪个变量作为因变量。然后把因变量用自变量的函数式表示。

1．设函数形式为F（x，y）=0，取变量y作为x的隐函数，即存在 $y=y(x)$ ，则可以把F（x，y）=0写成 $F(x,y(x))=0$ 的形式，两边对x求导，得到

$$\frac{d}{dx}F(x,y(x))=0$$

$$\frac{dy}{dx}=-\frac{F_x}{F_y}$$ ；

相应地，如果函数形式为F（x，y，z）=0，则有

$$\frac{\partial z}{\partial x}=-\frac{F_x}{F_z}$$ ；

$$\frac{\partial z}{\partial y}=-\frac{F_y}{F_z}$$ 。

（二）对于方程组，即

$$\begin{cases}F(x,y,z)=0\\G(x,y,z)=0\end{cases}$$

同样分别看成x的隐函数，得到

$$\begin{cases}F(x,y(x),z(x))=0\\G(x,y(x),z(x))=0\end{cases}$$

对每个方程的两边对x求导，解出方程组，得到

$$y'(x)=-\frac{\begin{vmatrix}F_x & F_z\\ G_x & G_z\end{vmatrix}}{\begin{vmatrix}F_y & F_z\\ G_y & G_z\end{vmatrix}};$$

$$z'(x)=-\frac{\begin{vmatrix}F_y & F_x\\ G_y & G_x\end{vmatrix}}{\begin{vmatrix}F_y & F_z\\ G_y & G_z\end{vmatrix}}$$

一般把上面的偏导数所组成的行列式称为雅可比式，并且简写为

$$\frac{\partial(F,G)}{\partial(y,z)}=\begin{vmatrix}F_y & F_z\\ G_y & G_z\end{vmatrix}$$

那么就可以简化上面的表达式。

?

高阶偏导数。

与单变量函数可以求高阶导数一样，多元函数同样可以求高阶偏导数。

可以想象，一个二元函数的偏导数是两个，而每一个偏导数又都是一个二元函数，因此再对这两个求偏导数，就最终得到四个二阶偏导数，即

对于 $z=f(x,y)$，存在 z_{xx}， z_{xy}， z_{yx}， z_{yy} 这样四个二阶偏导数。

注意到其中两个二阶偏导数 z_{xy} 和 z_{yx} 的差别只是在分别对x和y求导的先后顺序不同，对于它们的关系有定理：

如果 f_{xy} 和 f_{yx}在点（x，y）处都连续，则 $f_{xy}(x,y)=f_{yx}(x,y)$。

对于更高阶的偏导数，只要所涉及到的偏导数都连续，则存在相似的结论，即高阶的混合偏导数与对各个变量的求导先后顺序无关。

局部相对极值。

在单变量函数的极值定义当中，我们已经体会到，所谓极值只能是在它的某个邻域内有意义，这就是为什么需要称为局部相对极值的缘故，对于多变量函数也是相似的，多变量函数的局部相对极值的定义也是相似的，只是这里几个自变量被看成是一个点，而这个点实际上是由几个分量组成，这些分量正是函数的各个自变量，这样自变量值的某个邻域就是一个区域，在这个区域内所有其他的函数值都比极值大或者小。

如果一个点是极值点，首先，必须是临界点，也就是说，极值点的一个必要条件是这点必须是临界点，所谓临界点就是可微函数的梯度为0的点，这就是极值点的一阶必要条件。

进一步要分析判断一个临界点是否极值点，则还需要函数在这点的二阶偏导数，这就是下面的二阶充分条件：

设函数f在其临界点（a，b）的某个邻域内存在直到二阶的连续偏导数，下面我们定义一个简单记号

$$D=\begin{vmatrix}f_{xx}(x,y) & f_{xy}(x,y)\\ f_{yx}(x,y) & f_{yy}(x,y)\end{vmatrix},$$

那么有如下判别标准：

（1） $D>0, f_{xx}(a,b)>0$时，临界点（a，b）为函数的极小点；

（2） $D>0, f_{xx}(a,b)<0$时，临界点（a，b）为函数的极大点；

（3） $D<0$时，临界点（a，b）为函数的鞍点；

（4）D=0时，则不能判断临界点（a，b）的性质。

多元函数的最值，条件极值以及应用。

我们讨论极值的最终目的，还是要研究函数的最值问题，因为毕竟在实际问题当中，我们更经常要求解决的是最值问题。而最值问题是以极值问题为基础的，因为除了有可能在有界闭区域边界处取最值（称为条件极值）以外，最值肯定是属于极值的。

对于条件极值，我们有如下定义：

设函数f（x，y）在某点及其某个去心邻域内有定义，而对于这个去心邻域内的满足约束条件g（x，y）=0的所有点（x，y）处的函数值都小于或者大于在该点的函数值，则该点就是函数的局部相对极大值点或者是极小值点，相应的函数值为局部相对极大值或者是极小值。

运用这个定义的关键是约束条件的恰当表述。

至此我们就可以得到求最值的步骤：

（1） 计算函数的梯度，根据临界点的判别条件，求出所有的临界点；

（2） 剔除落在要求区域之外的临界点，然后根据二阶条件分析临界点；

（3） 分析有界闭区域的边界上的条件极值情况；

（4） 比较步骤（2）和步骤（3）所得到的结果，得到最终的最值。

拉格朗日乘数法。

在求解具有等式约束条件的条件极值问题时，一般并不是从约束等式解出一个变量，再代入目标函数，因为从约束等式解出一个变量往往并不简单，反而相当麻烦，因此，我们一般使用所谓的拉格朗日乘数法。

对于一个等式约束问题，即在约束条件g（x，y）=0下，求函数f（x，y）的最值的问题，可以从几何的角度，这样来考虑：把约束条件g（x，y）=0看成是XY平面上的一条曲线，而在XY平面上，可以给出函数f（x，y）的等值线，显然，所谓最值就一定是等值线与曲线g（x，y）=0相切的点。从几何的角度来求这样的切点，就可以得到拉格朗日乘数法。

首先切点的特征就是在这点上，曲线g（x，y）=0与函数f（x，y）的那条等值线具有相同的法线方向，即存在常数k，使得

$$\nabla f = -k\nabla g ,$$

把这个等式与约束条件联立求解，就能够得到条件极值。换一种看法，就是引入所谓拉格朗日函数：

$$L(x,y,k) = f(x,y) + kg(x,y) ,$$

其中k为常数，称为拉格朗日乘子，然后求函数L的临界点，就是求解下列方程组的解：

$$\begin{cases} \dfrac{\partial}{\partial x} L(x,y,k) = f_x(x,y) + k\, g_x(x,y) = 0 \\ \dfrac{\partial}{\partial y} L(x,y,k) = f_y(x,y) + k\, g_y(x,y) = 0 \\ \qquad \dfrac{\partial}{\partial k} L(x,y,k) = g(x,y) = 0 \end{cases}$$

这里实际上就是上面切点所满足的几何条件和约束等式的联立方程组，它的解就是可能的最值点。这种解法就是所谓的拉格朗日乘数法。

二，答疑解难。

关于多元函数具有的一些特有问题，我们必须极大地注意到，因此我们需要有心于把一元函数与多元函数的基本概念加以比较，有不同的地方，常常也就是我们容易犯错误的地方。例如，多元函数在某点的所有偏导数都存在，是否就一定在该点可微？

[答]：不一定。

多元函数在某点的所有偏导数都存在，并不能够保证函数在该点是连续的，而函数在该点可微的必要条件是在该点连续，因此必然同样也不能保证在该点可微。

一般地，二元函数在某点的这么一些性质：连续，可微，存在偏导数，偏导数连续的相互关系如何？

[答]：偏导数连续当然一定存在偏导数；

偏导数连续则一定可微；

可微则一定存在偏导数；

可微则一定连续；

以上定理都是充分条件，都不是必要条件，因此

连续不一定可微；

可微不一定偏导数连续；

偏导数存在不一定可微；

偏导数存在不一定就连续。