第 4 章　多元函数的微积分

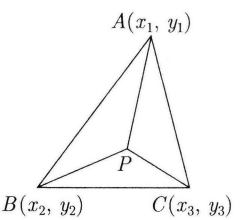
4.1　多元函数的微分

4.1.1　三角形的 Fermat 点

[**问题**]　设 $\triangle ABC$ 的三个内角都小于 $120°$. 要在其内部找一点 P, 使得 $PA+PB+PC$ 达到最小.

[**解答**]　如下图, 任意建立平面直角坐标系, 设 A,B,C 三点的位置向量分别为 $\boldsymbol{r}_1=(x_1,y_1)$, $\boldsymbol{r}_2=(x_2,y_2)$, $\boldsymbol{r}_3=(x_3,y_3)$, P 点的位置向量为 $\boldsymbol{r}=(x,y)$, 那么 $PA+PB+PC$ 的值为

$$L(\boldsymbol{r})=\sum_{j=1}^{3}||\boldsymbol{r}-\boldsymbol{r}_j||.$$



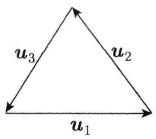
对 L 求偏导数, 有

$$\begin{aligned}\frac{\partial L}{\partial x}&=\sum_{j=1}^{3}\frac{\partial}{\partial x}\sqrt{(x-x_j)^2+(y-y_j)^2}\\&=\sum_{j=1}^{3}\frac{x-x_j}{\sqrt{(x-x_j)^2+(y-y_j)^2}}=\sum_{j=1}^{3}\frac{x-x_j}{||\boldsymbol{r}-\boldsymbol{r}_j||},\end{aligned}$$

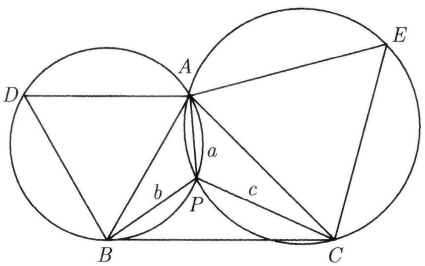
同理,

$$\frac{\partial L}{\partial y}=\sum_{j=1}^{3}\frac{y-y_j}{||\boldsymbol{r}-\boldsymbol{r}_j||}.$$

记 $\boldsymbol{u}_i = \dfrac{\boldsymbol{r}-\boldsymbol{r}_i}{||\boldsymbol{r}-\boldsymbol{r}_i||}$ $(i=1,2,3)$, 则 $||\boldsymbol{u}_1|| = ||\boldsymbol{u}_2|| = ||\boldsymbol{u}_3|| = 1$, 且由 $\dfrac{\partial L}{\partial x} = 0$ 及 $\dfrac{\partial L}{\partial y} = 0$ 得 $\boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{u}_3 = 0$. 于是, 向量 $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ 可首尾相接构成正三角形的三边, 如下图.



所以 $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ 两两之间的夹角都是 120°. 上述对 $\triangle ABC$ 的三条边的张角都是 120° 的点称为 $\triangle ABC$ 的 Fermat 点. 根据平面几何的基本知识, 当 $\triangle ABC$ 的三个内角都小于 120° 时, 分别以 AB, AC 为一边向外作正三角形 $\triangle ABD$ 和 $\triangle ACE$, 并作它们的外接圆, 这两圆必在 $\triangle ABC$ 内部交于一点, 此即为 Fermat 点, 如下图中的 P 点.

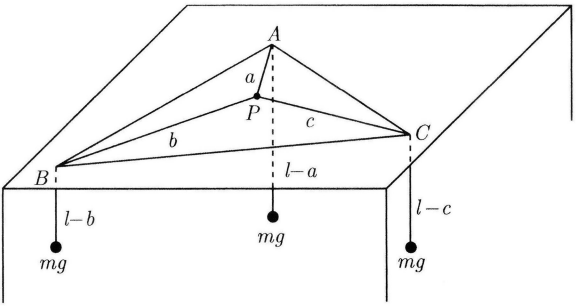


现在记 P 点的位置向量为 $\boldsymbol{r}_0$. 由 $\lim\limits_{\boldsymbol{r}\to\infty} L(\boldsymbol{r}) = +\infty$ 得, $\mathbf{R}^2$ 上的连续函数 $L(\boldsymbol{r})$ 必能取到最小值, 取到最小值的可能点必为驻点或偏导数不存在的点, 即 P, A, B, C. 记 $a = PA$, $b = PB$, $c = PC$, 则

$$\begin{aligned} L(\boldsymbol{r}_1) &= AB + AC = \sqrt{a^2+b^2+ab} + \sqrt{a^2+c^2+ac} \\ &> \left(\frac{a}{2}+b\right) + \left(\frac{a}{2}+c\right) = a+b+c = L(\boldsymbol{r}_0). \end{aligned}$$

对 $\boldsymbol{r}_2$, $\boldsymbol{r}_3$, 有类似的不等式, 即 $L(\boldsymbol{r}_0) < L(\boldsymbol{r}_j)$ $(j=1,2,3)$. 所以 L 在 Fermat 点达到最小值.

[**说明**]　下面来看物理 "证明". 如下图, 在一个光滑的桌面上, 在 $\triangle ABC$ 的三个顶点处开三个洞, 并假设洞处也是光滑的. 将三根长为 l 的无质量细绳的一端各连上一个质量为 m 的小球, 将细绳从洞中穿上桌面后将三根绳的另一端系在一起, 记该点为 P, 然后等系统达到平衡.



这时, 三个小球的总势能为

$$U = -mg(l-a) - mg(l-b) - mg(l-c) = -3mg + mg(a+b+c).$$

在系统平衡时, U 达到极小, 而这当且仅当 $a+b+c$ 达到极小. 由于三根细绳上的张力相等, 这三个力能首尾相接构成一个正三角形. 所以在 P 点, 任何两根细绳之间的夹角都是 $120°$, P 必是 $\triangle ABC$ 的 Fermat 点.

上述物理 "证明" 与前面给出的用偏导数的证明是一致的. 势能的负梯度就是力, 势能达到极小时其梯度为零, 也就意味着受力为零.

下面来看一般的结论. 设质点受的力仅同位置有关而同速度无关, 且是保守力①. 记 $\boldsymbol{r} = (x_1, x_2, x_3)$ 为坐标向量, 并设 $\boldsymbol{r} = 0$ 是平衡点. 由于力 $\boldsymbol{F} = (F_1, F_2, F_3)$ 是保守力, 存在势能 $U(\boldsymbol{r})$, 使得 $\boldsymbol{F}(\boldsymbol{r}) = -\nabla U(\boldsymbol{r})$. 原点是平衡点意味着 $\boldsymbol{F}(0) = 0$. 于是, 由 Taylor 公式,

$$\begin{aligned} F_i(\boldsymbol{r}) &= \sum_{j=1}^{3} \frac{\partial F_i}{\partial x_j}(0) x_j + o(||\boldsymbol{r}||) \\ &= -\sum_{j=1}^{3} \frac{\partial^2 U}{\partial x_i \partial x_j}(0) x_j + o(||\boldsymbol{r}||), \end{aligned}$$

从而

$$\boldsymbol{F}(\boldsymbol{r}) \cdot \boldsymbol{r} = -\sum_{i,j=1}^{3} \frac{\partial^2 U}{\partial x_i \partial x_j}(0) x_j x_j + o(||\boldsymbol{r}||^2). \tag{4.1}$$

以原点为球心、$||\boldsymbol{r}||$ 为半径作球面 $S_{||\boldsymbol{r}||}$. 如果 U 在原点的 Hesse (黑塞) 矩阵

$$\left(\frac{\partial^2 U}{\partial x_i \partial x_j}(0) \right)_{i,j=1,2,3}$$

① 当质点受力同速度有关时 (例如磁场力或 Coriolis 力), 这里的说明无效, 必须考虑包括速度在内的一个方程组的解的稳定性.

为正定, 则 U 在原点达到极小值. 这时, 由 (4.1) 式, 当一个静止质点在原点 (平衡点) 附近时, 它在该点的受力指向相应的球面 $S_{||r||}$ 内侧, 也就是说, 该质点将靠近原点. 这时原点是稳定平衡点. 反之, 如果 U 在原点的 Hesse 矩阵有一个负特征值, 则在对应的特征子空间中, U 在原点达到极大值. 这时原点是不稳定平衡点.

4.1.2　最小二乘法

1801 年, 天文学家 Piazzi (皮亚奇) 发现了一个移动的天体, 但其后该天体运动到太阳附近从而无法观察, 等到预计已离开太阳的位置时, 早就不知去向, 经许多天文学家的努力也未找到. 于是, 数学家 Gauss (高斯) 接手了此事, 他利用自己建立起来的最小二乘法, 靠 Piazzi 仅有的几个观察数据, 计算出了这个天体的轨道, 从而使天文学家重新找到了这个天体, 这就是小行星带中最大的小行星——谷神星.

现在最小二乘法用得比较多的地方是拟合直线, 事实上它可以用来拟合任意给定形状的曲线. 对于直线的拟合, 设给出 n 个点 (x_i, y_i) $(i = 1, 2, \cdots, n)$, 要找参数 a, b, 使得 $\sum\limits_{i=1}^{n}(ax_i + b - y_i)^2$ 达到最小, 这时 $y = ax + b$ 是同已给的点 "最接近" 的直线. 接下来只需简单地用偏导数求极值即可得到 a, b. 需要注意的是, 这里的函数 $\sum\limits_{i=1}^{n}(ax_i + b - y_i)^2$ 是人为定义的, 也就是说, "最接近" 是人为规定的, 也可以用例如 $\sum\limits_{i=1}^{n}|ax_i + b - y_i|$. 事实上, 直观上后者更合理些, 因为它等价于各点到直线的距离之和最小. 通常要求平方和最小纯粹是为了求导方便.

对于最小二乘法, 重要的是了解使什么最小以及如何求最小值, 仅记住直线拟合公式或能用软件计算对数学学习是毫无意义的. 下面做一个和 Gauss 计算谷神星轨道类似的工作.

[**问题**]　给定平面上的点 (x_i, y_i) $(n = 1, 2, \cdots, n)$, 求一个焦点在原点的椭圆, 使之同这 n 个点最接近.

[**解答**]　由于椭圆的焦点在原点, 我们使用极坐标下的椭圆方程

$$r = \frac{ep}{1 - e\cos(\theta - \theta_0)}.$$

记 $\lambda = ep$, $\alpha = e\cos\theta_0$, $\beta = e\sin\theta_0$, 则方程成为

$$r - \alpha x - \beta y = \lambda,$$

其中 $x = r\cos\theta$, $y = r\sin\theta$. 于是, 作

$$L = \sum_{i=1}^{n}(\alpha x_i + \beta y_i + \lambda - r_i)^2,$$

其中 $r_i = \sqrt{x_i^2 + y_i^2}$. 要确定常数 λ, α, β, 使得 L 达到最小.

对两个 n 维向量 $(u_1, \cdots, u_n)$ 和 $(v_1, \cdots, v_n)$, 记 $S_{uv} = \sum\limits_{j=1}^{n} u_j v_j$, $S_u = \sum\limits_{j=1}^{n} u_j$. 由 $\dfrac{\partial L}{\partial \lambda} = 0$, $\dfrac{\partial L}{\partial \alpha} = 0$, $\dfrac{\partial L}{\partial \beta} = 0$ 得到关于 (λ, α, β) 的线性代数方程组

$$\begin{aligned} &S_x\alpha + S_y\beta + n\lambda = S_r, \\ &S_{xx}\alpha + S_{xy}\beta + S_x\lambda = S_{xr}, \\ &S_{yx}\alpha + S_{yy}\beta + S_y\lambda = S_{yr}. \end{aligned} \tag{4.2}$$

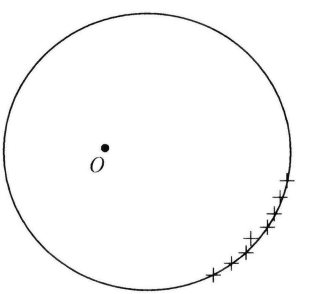
解此线性代数方程组即得 (λ, α, β).

[**说明**] 例如, 现有 8 个点,

$$\begin{aligned} &(12.73, -16.82), \quad (15.11, -15.26), \quad (17.5, -13.82), \quad (17.68, -11.93), \\ &(19.87, -10.45), \quad (20.77, -8.68), \quad (21.51, -6.51), \quad (22.43, -4.30), \end{aligned}$$

要求一个焦点在原点的椭圆, 使之同这 8 个点最接近.

对已知的数据, 解方程组 (4.2) 得 $\lambda = 17.80$, $\alpha = 0.21$, $\beta = -0.03$. 所求得的椭圆与已知的 8 个点见下图, 其中 O 表示原点.



由于给定的 8 个点偏于一侧, 所得的椭圆必然会有较大的误差, 最小二乘法只是将误差尽量减少. 如果这 8 个点在椭圆上比较均匀地分布, 结果就会准确得多. 当年关于谷神星的数据比这个例子中的更集中, 并且由于地球在运动, 谷神星相对地球的轨道远比椭圆复杂, 这就是在 Gauss 使用最小二乘法之前这么多天文学家无法确定它的轨道的原因.