非线性最佳平方拟合

前面介绍的数据拟合最小二乘法属于线性拟合

已知数据点 $\{x_i, f(x_i)=y_i\}_{i=0}^{n}$

$\varphi_0, \varphi_1, \cdots \varphi_m$ 为线性无关的基函数

取 $P(x)=a_0\varphi_0(x)+\cdots+a_m\varphi_m(x)$ 线性叠加

使 $Q(a_0, a_1, \cdots a_m)=\sum_{i=0}^{n}\left[P(x_i)-f(x_i)\right]^2$ 最小

实际问题：

（1）拟合函数不能表示为参量的线性叠加形式，如：

$$P(x)=a_0+a_1\sin(a_2x)$$

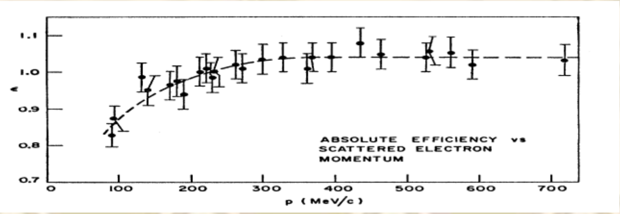
$$P(x)=a_0\exp[-\frac{1}{2}(\frac{x-a_1}{a_2})]+a_3+a_4x+a_5x^2$$

（2）拟合函数与参量的关系无法用解析表达式给出，仅知

$P(x,a_0,a_2,\cdots,a_n)$ 可由理论模型数值计算得离散点值。

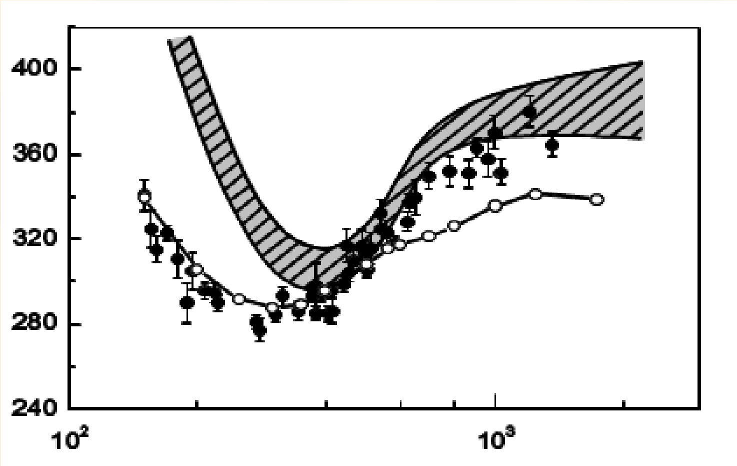
拟合的目的是找参量 $\{a_k\}_{k=0}^{n}$ 使计算结果逼近实验结果。

前面的线性拟合方法无法实现！



例如：

测量数据为 $\{x_i,\ f(x_i),\ \sigma_i\}_{i=0}^{N}$



由理论模型方法给出其满足运动方程：

$$\nabla^2 y(x) + F\{a_1,\ a_2,\ \cdots,\ a_M,\ x,\ y(x)\} = 0$$

对给定的参量 $\{a_k\}_{k=1}^{M}$ 可通过数值方法计算得出：$\{x_i,\ y(x_i)\}$

目的：寻找 $\{a_k^*\}_{k=1}^{M}$ 使得 $\{y(a_1^*\cdots a_M^*,\ x_i)\}$ 很好地逼近实验。

非线性最佳平方拟合方法基本原理相同：

实验数据点 $\{x_i, f(x_i)=y_i, \sigma_i\}_{i=0}^{N}$

找 $P(x)$, $\left\{a_k\right\}_{k=1}^{n}$ 或 $P(x, a_1, a_2, \cdots, a_n)$

使 $Q(a_1, a_2, \cdots a_n) = \sum_{i=0}^{N} \frac{1}{\sigma_i^2}\left[y_i - P(x_i, a_1, a_2, \cdots, a_n)\right]^2$ 最小

搜索法： $Q(a_1, a_2, \cdots a_n)$ 看成 $(a_1, a_2, \cdots, a_n)$ 的连续函数。

在这n维空间，由一组初值 $\left\{a_k^0\right\}_{k=1}^{n}$ 出发

$\left\{a_k^0\right\}_{k=1}^{n} \Rightarrow \left\{a_k^1\right\}_{k=1}^{n} \Rightarrow \cdots \Rightarrow \left\{a_k^*\right\}_{k=1}^{n}$ 达到 $Q(a_1^*, a_2^*, \cdots, a_n^*)$ 最小

问题： $\left\{a_k^0\right\}_{k=1}^{n} \Rightarrow \left\{a_k^1 = a_k^0 + \delta a_k^0\right\}_{k=1}^{n}$ 如何确定增量 $\left\{\delta a_k^0\right\}_{k=1}^{n}$

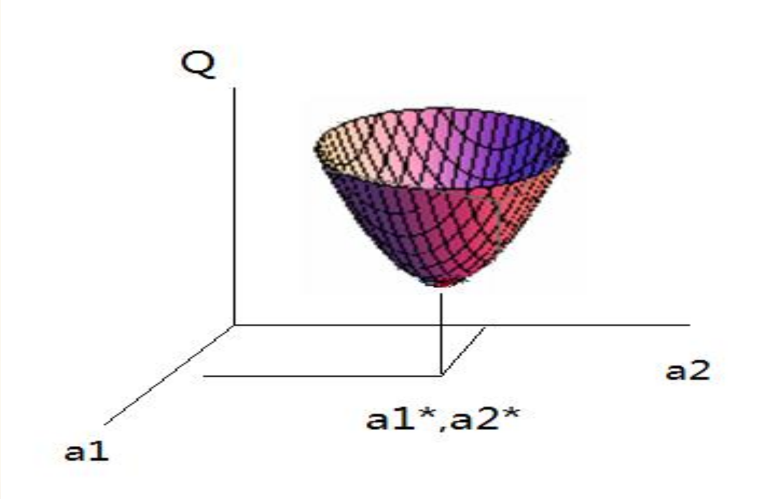
1、梯度搜索法(Gradient Search)(最速下降法)

(1) 先看2参数的情况

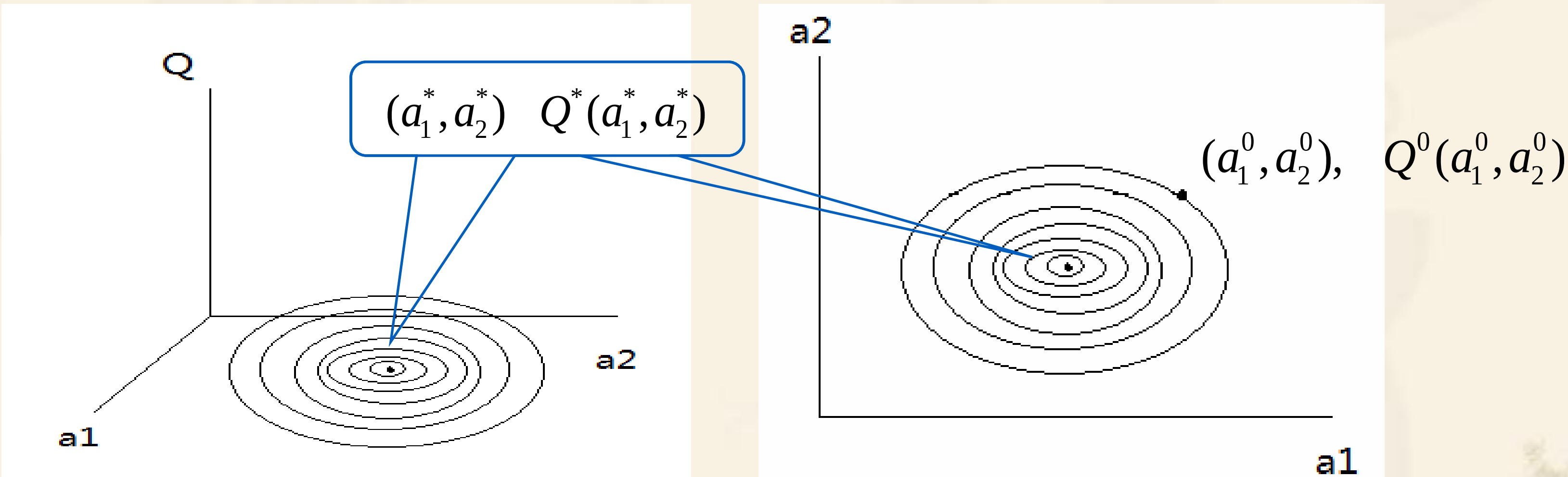
实验数据点 $\{x_i, f(x_i) = y_i, \sigma_i\}_{i=0}^{N}$

找 $P(x, a_1, a_2)$ 使 $Q(a_1, a_2) = \sum_{i=0}^{N} \frac{1}{\sigma_i^2}\left[y_i - P(x_i, a_1, a_2)\right]^2$ 最小

$Q(a_1, a_2)$ 为三维空间的曲面，最低点 $Q^*(a_1^*, a_2^*)$



用一系列平行于(a1, a2)的平面切割曲面，投影到(a1, a2)平面，得到一系列等高线，每一曲线 Q相同，不同的线表不同的 Q，最小点为 $Q^*(a_1^*, a_2^*)$



(a) 取初始点 (a_1^0, a_2^0), $Q^0(a_1^0, a_2^0)$ 为相应的等高线

找下一点 $(a_1^1 = a_1^0 + \delta a_1^0, a_2^1 = a_2^0 + \delta a_2^0)$, 使$Q^1(a_1^1, a_2^1)$ 为初始点附近最小

(b)梯度方向Q上升最快

$$\nabla Q=\frac{\partial Q}{\partial a_1}\hat{a}_1+\frac{\partial Q}{\partial a_2}\hat{a}_2$$ 梯度沿各参量轴的分量为：

$$\chi_k=\frac{\partial Q}{\partial a_k}/\sqrt{\sum_i\left(\frac{\partial Q}{\partial a_\mathrm{i}}\right)^2}$$

取 $\delta a_k=-\Delta a_k\chi_k$，$\Delta a_k$ 为选定的步长（可自由调节）。

则 $Q^1(a_1^1,a_2^1)$ 为初始点附近**Q**最小

同样可得： $Q^0(a_1^0,a_2^0)>Q^1(a_1^1,a_2^1)>Q^2(a_1^2,a_2^2)>\cdots>Q^*(a_1^*,a_2^*)$

注： $$\frac{\partial Q}{\partial a_1}\approx\frac{Q(a_1+\Delta a_1,a_2)-Q(a_1,a_2)}{\Delta a_1}$$

(2) 多参数的情况

实验数据点 $\{x_i, f(x_i)=y_i, \sigma_i\}_{i=0}^{N}$

使 $Q(a_1,a_2,\cdots,a_n)=\sum_{i=0}^{N}\frac{1}{\sigma_i^2}[y_i-P(x_i,a_1,a_2,\cdots,a_n)]^2$ 最小

$Q(a_1,a_2,\cdots,a_n)$ 为n维空间的曲面，最低点 $Q^*(a_1^*,a_2^*,\cdots,a_n^*)$

同样有一系列的等高线，每一曲线 Q相同，不同的线表不同的 Q ，最小点为 $Q^*(a_1^*,a_2^*,\cdots,a_n^*)$

$\nabla Q=\sum_k\left(\frac{\partial Q}{\partial a_k}\hat{a}_k\right)$ 梯度沿各参量轴的分量为:

$$\chi_k=\frac{\partial Q}{\partial a_k}\Big/\sqrt{\sum_i\left(\frac{\partial Q}{\partial a_i}\right)^2}$$

取 $\delta a_k = -\Delta a_k \chi_k$，$\Delta a_k$ 为选定的步长（可自由调节）。

则 $Q^1(a_1^1, a_2^1, \cdots, a_n^1)$ 为初始点附近Q最小

同样可得：

$$Q^0(a_1^0, a_2^0, \cdots, a_n^0) > Q^1(a_1^1, a_2^1, \cdots, a_n^1) > \cdots > Q^*(a_1^*, a_2^*, \cdots, a_n^*)$$

注：$$\frac{\partial Q}{\partial a_k} \approx \frac{Q(a_1, \cdots, a_k + \Delta a_k, \cdots, a_n) - Q(a_1, a_2, \cdots, a_n)}{\Delta a_k}$$

非线性最佳平方拟合梯度搜索法步骤：

实验数据点 $\{x_i, f(x_i) = y_i, \sigma_i\}_{i=0}^{N}$

找 $P(x)$, $\{a_k\}_{k=1}^{n}$ 或 $P(x, a_1, a_2, \cdots, a_n)$

(1) 任取参数初值 $\left\{a_k^0\right\}_{k=1}^{n}$ 计算 $Q^0(a_1^0, \cdots, a_n^0)$

(2) 取各参数的步长(可自由调节) $\Delta a_k^0, k = 1, \cdots, n$

计算梯度沿各参量轴的分量： $\chi_k = \dfrac{\partial Q}{\partial a_k} \Big/ \sqrt{\sum_i \left(\dfrac{\partial Q}{\partial a_i}\right)^2}$

得 $\delta a_k^0 = -\Delta a_k \chi_k \Rightarrow \left\{a_k^1 = a_k^0 + \delta a_k^0\right\}_{k=1}^{n} \Rightarrow Q^1(a_1^1, a_2^1, \cdots, a_n^1)$

(3) 再同样(1)(2)得 $Q^2, \cdots, Q^*(a_1^*, a_2^*, \cdots, a_n^*)$ 最小

2、网格搜索法(Grid Search)

（略）

3、抛物线外推法(Parabolic Extrapolation)

（略）