4月3日作业

1

1. In the context of data fitting, a few nonlinear models are called intrinsically linear because they can be converted to linear models by certain transformations. Show that the following nonlinear models are intrinsically linear (under mild assumptions) by finding appropriate transformations.
(1) $y = a\exp(bx + cx^2)$.
(2) $y = 1/\big(1 + \exp(a + bx)\big)$.
(3) $y = ax/(b + \sqrt{x}\,)^2$.

解：下面都假设有n个数据点 $(x_1, y_1) \cdots (x_n, y_n)$

(1) $y = a\exp(bx + cx^2)$

$\Rightarrow \ln y = \ln a + bx + cx^2$

令 $\tilde{a} = \ln a$，则：

$$\ln y = \tilde{a} + bx + cx^2$$

解以下最小二乘问题：

$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} \tilde{a} \\ b \\ c \end{bmatrix} = \begin{bmatrix} \ln y_1 \\ \ln y_2 \\ \vdots \\ \ln y_n \end{bmatrix}$$

得到最小二乘解 $\tilde{a}, b, c$ 后令 $a = e^{\tilde{a}}$ 即可

(2) $y = 1/(1 + \exp(a + bx))$

$$\Rightarrow \frac{1}{y} = 1+\exp(a+bx)$$

$$\ln\left(\frac{1}{y}-1\right) = a+bx$$

则 $\ln\left(\frac{1}{y}-1\right)$ 关于 x 是线性的

(3) $y = ax/(b+\sqrt{x})^2$

$$\Rightarrow \frac{ax}{y} = (b+\sqrt{x})^2$$

$$\Rightarrow \sqrt{a}\cdot\sqrt{\frac{x}{y}} = b+\sqrt{x}$$

$$\Rightarrow \sqrt{\frac{x}{y}} = \frac{b}{\sqrt{a}} + \frac{\sqrt{x}}{\sqrt{a}}$$

则 $\sqrt{\frac{x}{y}}$ 关于 $\sqrt{x}$ 是线性的 $\square$

2

2. Chebyshev polynomials can be expressed as $T_n(x) = \cos(n \arccos x)$ for $x \in [-1,1]$. Can you find an explicit expression of $T_n(x)$ for $x \in \mathbb{R} \setminus [-1,1]$ in terms of algebraic functions?

解：Chebyshev 多项式有如下递推关系：

$$T_n = 2x T_{n-1} - T_{n-2} \qquad x \in \mathbb{R} \setminus [-1, 1]$$

解特征方程: $\lambda^2 = 2x\lambda - 1$

$$\Rightarrow \lambda_1 = \frac{2x + \sqrt{4x^2 - 4}}{2} = x + \sqrt{x^2 - 1}$$

$$\lambda_2 = \frac{2x - \sqrt{4x^2 - 4}}{2} = x - \sqrt{x^2 - 1}$$

设 $T_n = \mu_1 \cdot \lambda_1^n + \mu_2 \cdot \lambda_2^n$

由:

$$\begin{cases} T_0 = \mu_1 + \mu_2 = 1 \\ T_1 = \mu_1 \cdot \lambda_1 + \mu_2 \lambda_2 = x \end{cases}$$

$$\Rightarrow \mu_1 = \mu_2 = \frac{1}{2}$$

$$\Rightarrow T_n = \frac{1}{2}\left(x + \sqrt{x^2 - 1}\right)^n + \frac{1}{2}\left(x - \sqrt{x^2 - 1}\right)^n$$

$$x \in \mathbb{R} \setminus [-1, 1]$$

$\square$

3

3. Find
$$\min_{a,b,c\in\mathbb{R}} \int_0^{\pi/2} |\sin x - ax^2 - bx - c|^2 \, dx$$
without programming.

解：定义函数空间上的内积 $\langle F, G\rangle = \int_0^{\frac{\pi}{2}} F(x)\cdot G(x)\cdot dx$

则原问题变为：

$$\min_{\phi\in S} \langle F-\phi^T d, F-\phi^T d\rangle, \quad d=(c,a,b)^T$$

其中 $\phi\cdot d = ax^2+bx+c$，$S=\mathrm{span}\{1, x, x^2\}$，$F=\sin x$

$\langle F-\phi d, F-\phi d\rangle$

$= \langle F, F\rangle - 2\langle F, \phi\rangle^T d + d^T\langle \phi, \phi\rangle d$

其中：

$\langle F, F\rangle = \int_0^{\pi/2} (\sin x)^2 dx = \frac{\pi}{4}$

$\langle F, \phi\rangle = (\langle \sin x, 1\rangle, \langle \sin x, x\rangle, \langle \sin x, x^2\rangle)^T$

$= (1, 1, \pi-2)^T$

$$\langle \phi, \phi\rangle = \begin{bmatrix} \langle 1,1\rangle & \langle 1,x\rangle & \langle 1,x^2\rangle \\ \langle x,1\rangle & \langle x,x\rangle & \langle x,x^2\rangle \\ \langle x^2,1\rangle & \langle x^2,x\rangle & \langle x^2,x^2\rangle \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\pi}{2} & \frac{1}{2}\left(\frac{\pi}{2}\right)^2 & \frac{1}{3}\left(\frac{\pi}{2}\right)^3 \\ \frac{1}{2}\left(\frac{\pi}{2}\right)^2 & \frac{1}{3}\left(\frac{\pi}{2}\right)^3 & \frac{1}{4}\left(\frac{\pi}{2}\right)^4 \\ \frac{1}{3}\left(\frac{\pi}{2}\right)^3 & \frac{1}{4}\left(\frac{\pi}{2}\right)^4 & \frac{1}{5}\left(\frac{\pi}{2}\right)^5 \end{bmatrix}$$

$$= D \cdot H \cdot D$$

其中 H 是 3 阶 Hilbert 矩阵

而 $D = \begin{bmatrix} \left(\frac{\pi}{2}\right)^{1/2} & & \\ & \left(\frac{\pi}{2}\right)^{3/2} & \\ & & \left(\frac{\pi}{2}\right)^{5/2} \end{bmatrix}$

$$\Rightarrow \langle \phi, \phi \rangle^{-1} = D^{-1} H^{-1} D^{-1}$$

其中 $(H^{-1})_{ij} = \dfrac{(-1)^{i+j}(n+i-1)!(n+j-1)!}{((i-1)!(j-1)!)^2(n-i)!(n-j)!(i+j-1)}$ （从书上搬过来的）

当 $\tilde{d} = -\langle \phi, \phi \rangle^{-1} \cdot \langle F, \phi \rangle$ 时有最小值

那么：

$$\min \langle F - \phi^T \tilde{d}, F - \phi^T \tilde{d} \rangle$$

$$= \langle F, F \rangle - \langle F, \phi \rangle^T \langle \phi, \phi \rangle^{-1} \langle F, \phi \rangle$$

$$= \frac{\pi}{4} - (1, 1, \pi - 2) D^{-1} \cdot H^{-1} \cdot D^{-1} (1, 1, \tfrac{\pi}{2})^T$$

$$= \frac{\pi}{4} - \left(\left(\frac{\pi}{2}\right)^{-1/2}, \left(\frac{\pi}{2}\right)^{-3/2}, (\pi-2)\cdot\left(\frac{\pi}{2}\right)^{-5/2} \right) \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix} \begin{pmatrix} \left(\frac{\pi}{2}\right)^{-1/2} \\ \left(\frac{\pi}{2}\right)^{-3/2} \\ (\pi-2)\left(\frac{\pi}{2}\right)^{-5/2} \end{pmatrix}$$

$$= \frac{\pi}{4} - \Big(9\cdot\left(\frac{\pi}{2}\right)^{-1} - 72\cdot\left(\frac{\pi}{2}\right)^{-2} + 60(\pi-2)\left(\frac{\pi}{2}\right)^{-3} + 192\left(\frac{\pi}{2}\right)^{-3} - 360(\pi-2)\cdot\left(\frac{\pi}{2}\right)^{-4} + 180(\pi-2)^2\left(\frac{\pi}{2}\right)^{-5} \Big)$$

$$\approx 1.1036\times 10^{-4}$$

(已经算到崩溃了!) □

4

4. Suppose that there is a bimodal function of the form

$$y = \alpha_1 g\big(\beta_1(x-\gamma_1)\big) + \alpha_2 g\big(\beta_2(x-\gamma_2)\big),$$

where $g(x) = \exp(-x^2)$. Can you find out the parameters from the following *noisy* data set sampled from this model?

解: $f(x) = \alpha_1 \cdot e^{-(\beta_1(x-\gamma_1))^2} + \alpha_2 e^{-(\beta_2(x-\gamma_2))^2}$

记 $r = \begin{bmatrix} y_1 - f(x_1) \\ \vdots \\ y_n - f(x_n) \end{bmatrix}$

在 $c^{(k)} = (\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2)$ 附近做线性化:

$$r^{(k+1)} \approx \left(J^{(k)}\right)\left(c^{(k+1)} - c^{(k)}\right) + r^{(k)}$$

希望 $r^{(k+1)}$ 尽量小，则取

$$c^{(k+1)} = c^{(k)} - \left(J^{(k)}\right)^{+} r^{(k)}$$

而 $f(x) = \alpha_1 \cdot e^{-(\beta_1(x-\gamma_1))^2} + \alpha_2 e^{-(\beta_2(x-\gamma_2))^2}$

$$\frac{\partial f}{\partial \alpha_1} = e^{-(\beta_1(x-\gamma_1))^2}$$

$$\frac{\partial f}{\partial \alpha_2} = e^{-(\beta_2(x-\gamma_2))^2}$$

$$\frac{\partial f}{\partial \beta_1} = \alpha_1 \cdot e^{-(\beta_1(x-\gamma_1))^2} \cdot (x-\gamma_1)^2 \cdot (-2\beta_1)$$

$$\frac{\partial f}{\partial \beta_2} = \alpha_2 \cdot e^{-(\beta_2(x-\gamma_2))^2} \cdot (x-\gamma_2)^2 (-2\beta_2)$$

$$\frac{\partial f}{\partial \gamma_1} = \alpha_1 \cdot e^{-\beta_1(x-\gamma_1)^2} \cdot (-\beta_1^2) \cdot (2(\gamma_1 - x))$$

$$\frac{\partial f}{\partial \gamma_2} = \alpha_2\, e^{-\beta_2(x-\gamma_2)^2} (-\beta_2^2)\, (2(\gamma_2 - x))$$

$$J^{(k)} = \begin{bmatrix} \frac{\partial f(x_1)}{\partial \alpha_1} & \frac{\partial f(x_1)}{\partial \alpha_2} & \frac{\partial f(x_1)}{\partial \beta_1} & \frac{\partial f(x_1)}{\partial \beta_2} & \frac{\partial f(x_1)}{\partial \gamma_1} & \frac{\partial f(x_1)}{\partial \gamma_2} \\ \vdots & & & & & \\ \frac{\partial f(x_n)}{\partial \alpha_1} & \cdots & & & & \frac{\partial f(x_n)}{\partial \gamma_2} \end{bmatrix}$$

对数据点进行观察后，使用下述参数作为初始估计：

$\alpha_1^{(0)}=2 \quad \alpha_2^{(0)}=1 \quad \beta_1^{(0)}=2, \ \beta_2^{(0)}=1, \ \gamma_1^{(0)}=-2.5, \ \gamma_2^{(0)}=1.5$

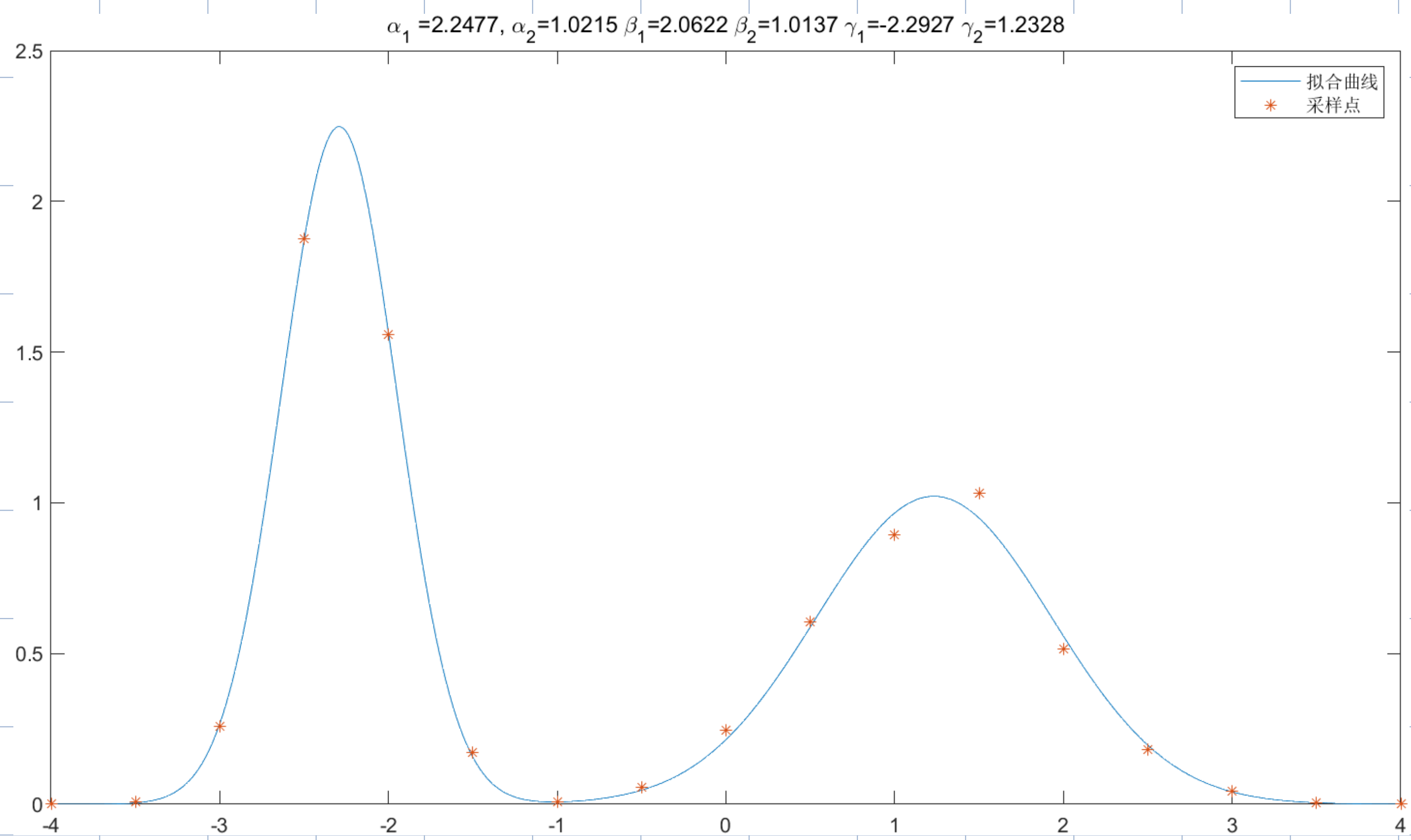
最终收敛到：

$\alpha_1 = 2.2477 \qquad \alpha_2 = 1.0215$

$\beta_1 = 2.0622 \qquad \beta_2 = 1.0137$

$\gamma_1 = -2.2927 \qquad \gamma_2 = 1.2328$

拟合曲线效果为：



$\square$

Q：有个奇怪的之处，既然在 $c^{(k)}$ 处的 Jacobi 阵是梯度，那么当 $\|r\|_2$ 取极值时梯度应该为零，但最终 $\|J\| \approx 5$，为什么呢？

5

5. (optional) In the homework on March 20, you have been asked to interpolate temperature in human body using a cubic spline. In fact, interpolation is not a good idea because the data are noisy. Try to fit the data using a periodic cubic spline with equispaced nodes $\{1, 4, 7, \ldots, 19, 22, 25\}$. Plot your solution for a two-day-period, and compare it with the interpolating spline.

解：设等距节点分别为 $x_1, x_2, \cdots, x_{n+1}$

在 $[x_i, x_{i+1}]$ 上的三次样条函数设为

$$f_i(x) = a_i x^3 + b_i \cdot x^2 + c_i x + d_i$$

并且 f_i 与 f_{i+1} 满足：

$$\begin{cases} f_i(x_{i+1}) = f_{i+1}(x_{i+1}) & \text{函数值连续} \\ f_i'(x_{i+1}) = f_{i+1}'(x_{i+1}) & \text{一阶导连续} \\ f_i''(x_{i+1}) = f_{i+1}''(x_{i+1}) & \text{2阶导连续} \end{cases}$$

也即：

$$\begin{bmatrix} x_{i+1}^3 & x_{i+1}^2 & x_{i+1} & 1 \\ 3x_{i+1}^2 & 2x_{i+1} & 1 & 0 \\ 6x_{i+1} & 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix} = \begin{bmatrix} x_{i+1}^3 & x_{i+1}^2 & x_{i+1} & 1 \\ 3x_{i+1}^2 & 2x_{i+1} & 1 & 0 \\ 6x_{i+1} & 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{i+1} \\ b_{i+1} \\ c_{i+1} \\ d_{i+1} \end{bmatrix}$$

而对于 f_n 与 f_1 而言，其关系为

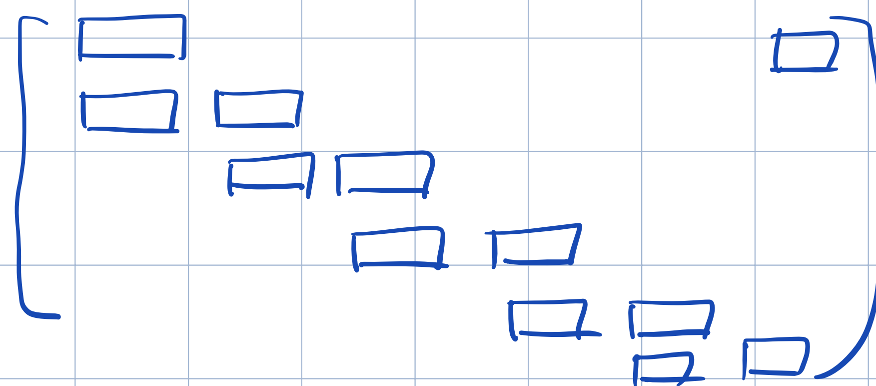
$$\begin{bmatrix} x_{n+1}^3 & x_{n+1}^2 & x_{n+1} & 1 \\ 3x_{n+1}^2 & 2x_{n+1} & 1 & 0 \\ 6x_{n+1} & 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_n \\ b_n \\ c_n \\ d_n \end{bmatrix} = \begin{bmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ 3x_1^2 & 2x_1 & 1 & 0 \\ 6x_1 & 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \end{bmatrix}$$

(注：也即还剩下 n 个自由度)

若记 $g = (a_1, b_1, c_1, d_1, \cdots, a_n, b_n, c_n, d_n)^T$

则 g 满足约束 $Ag=0$，其中 A 是一个长方形矩阵：

$A =$



$3n \times 4n$

再设处于 $[x_i, x_{i+1})$ 之间的数据点为：

$$(t_1^{(i)}, y_1^{(i)}) \cdots (t_{k_i}^{(i)}, y_{k_i}^{(i)})$$

记：

$$T_i = \begin{bmatrix} t_1^{(i)3} & t_1^{(i)2} & t_1^{(i)} & 1 \\ \vdots & & & \vdots \\ t_{k_i}^{(i)3} & t_{k_i}^{(i)2} & t_{k_i}^{(i)} & 1 \end{bmatrix}_{k_i \times 4}$$

$$Y = (y_1^{(1)} \cdots y_{k_1}^{(1)}, \cdots, y_1^{(n)} \cdots y_{k_n}^{(n)})^T$$

$$T = \begin{bmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_n \end{bmatrix}_{\sum_{i=1}^{n} k_i \times 4n}$$

则有： $T\cdot g \approx Y$

残差 $r = Y - T\cdot g$

则问题变成： $\min\ r^T r$

$\text{s.t.}\ A\cdot g = 0$ (江老师最优化课上学的)

这显然是一个凸问题，可以用软件包求解

(也有解析解，但我懒得算)

最后拟合结果如下：