3. $(x_i, y_i)_{i=1}^{n}$

最小二乘 $y = a + bx$，$y = \beta_0 + \beta_1 x + \beta_2 x^2$ 分别求解.

二次式残差平方和 $\le$ 一次式残差平方和 (*)

对一次回归 $Q = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = l_{yy} - \frac{l_{xy}^2}{l_{xx}}$，$\hat{b} = \frac{l_{xy}}{l_{xx}} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$，$\hat{a} = \bar{y} - \bar{x}\hat{b}$

对二次回归，$X = \begin{pmatrix} 1 & x_1 & x_1^2 \\ & \vdots & \\ 1 & x_n & x_n^2 \end{pmatrix}$ 满秩

$X^T X \beta = X^T Y$

$$X^T X = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \\ x_1^2 & \cdots & x_n^2 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & & \\ 1 & x_n & x_n^2 \end{pmatrix} = \begin{pmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{pmatrix}$$

$$\hat{\beta} = \begin{pmatrix} n & \sum x_i & \sum x_i^2 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum x_i^2 y_i \end{pmatrix}$$

对于(*)，$Q(\beta_0, \beta_1, \beta_2) = \inf_{\alpha_0, \alpha_1, \alpha_2} \sum_{i=1}^{n} (y_i - \alpha_0 - \alpha_1 x_i - \alpha_2 x_i^2)$

$$\le \inf_{\alpha_0, \alpha_1} \sum_{i=1}^{n} (y_i - \alpha_0 - \alpha_1 x_i)^2 = Q(a, b).$$

4.

x	18	20	22	24	26	28	30
y	26.86	28.35	28.75	28.87	29.75	30	30.36

$y = \beta_0 + \beta_1 x + e$

求 $\hat{\beta}_0$，$\hat{\beta}_1$，画图

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = 0.26429$$

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 22.6486$

绘图见后面页面

5. $y_1 = a + e_1$

$y_2 = 2a - b + e_2$

$y_3 = a + 2b + e_3$

e_1, e_2, e_3 独立, $Ee_i = 0$, $Var(e_i) = \sigma^2$ 求 $\hat{a}, \hat{b}$

此为多元回归问题 $X = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}$ $rank(X) = 2$

$X^TX = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 5 \end{pmatrix}$ $X^Ty = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} y_1 + 2y_2 + y_3 \\ -y_2 + 2y_3 \end{pmatrix}$

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = (X^TX)^{-1} X^Ty = \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} y_1 + 2y_2 + y_3 \\ -y_2 + 2y_3 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{6}(y_1 + 2y_2 + y_3) \\ \frac{1}{5}(-y_2 + 2y_3) \end{pmatrix}$$

$\hat{a} = \frac{1}{6}(y_1 + 2y_2 + y_3)$, $\hat{b} = \frac{1}{5}(-y_2 + 2y_3)$

6. $y_i = \beta_0 + \beta_1 x_i + \beta_2(3x_i^2 - 2) + e_i$, $i = 1, 2, 3$ 求 $\hat{\beta_0}, \hat{\beta_1}, \hat{\beta_2}$

e_i 两两不相关, $Ee_i = 0$, $Var(e_i) = \sigma^2$

$X = \begin{pmatrix} 1 & x_1 & 3x_1^2 - 2 \\ 1 & x_2 & 3x_2^2 - 2 \\ 1 & x_3 & 3x_3^2 - 2 \end{pmatrix}$ 满秩

$$X^TX = \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ 3x_1^2 - 2 & 3x_2^2 - 2 & 3x_3^2 - 2 \end{pmatrix} \begin{pmatrix} 1 & x_1 & 3x_1^2 - 2 \\ 1 & x_2 & 3x_2^2 - 2 \\ 1 & x_3 & 3x_3^2 - 2 \end{pmatrix} = \begin{pmatrix} 3 & \sum x_i & 3\sum x_i^2 - 6 \\ \sum x_i & \sum x_i^2 & 3\sum x_i^3 - 2\sum x_i \\ 3\sum x_i^2 - 6 & 3\sum x_i^3 - 2\sum x_i & \sum(3x_i^2 - 2)^2 \end{pmatrix}$$

$$X^Ty = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum(3x_i^2 y_i - 2y_i) \end{pmatrix}$$

$$\hat{\beta} = \begin{pmatrix} \hat{\beta_1} \\ \hat{\beta_2} \\ \hat{\beta_3} \end{pmatrix} = \begin{pmatrix} 3 & \sum x_i & 3\sum x_i^2 - 6 \\ \sum x_i & \sum x_i^2 & 3\sum x_i^3 - 2\sum x_i \\ 3\sum x_i^2 - 6 & 3\sum x_i^3 - 2\sum x_i & \sum(3x_i^2 - 2)^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \sum(3x_i^2 y_i - 2y_i) \end{pmatrix}$$

7\.

x	2	4	6	8	10
y	64	138	205	285	360

$y=\beta_0+\beta_1 x+e \quad e\sim N(0,\sigma^2)$

求出 $\hat\beta_0, \hat\beta_1$ $\alpha=0.05$ 检验 $\beta_1=0$

$\hat\beta_1=\frac{\sum_{i=1}^n (x_i-\bar x)(y_i-\bar y)}{\sum_{i=1}^n (x_i-\bar x)^2}=36.95$

$\hat\beta_0=\bar y-\hat\beta_1\bar x=-11.3$

$F=\frac{U}{Q/(n-2)}\sim F(1,n-2)$

$W=\{F>\lambda\} \quad P(W)=\alpha$

$U=\hat b\, l_{xy}=54612$

$Q=l_{yy}-U=37$

$F=4416$

$\lambda=10$

拒绝假设，呈线性关系

$\sum x_i y_i-n\bar x\bar y-n\bar y\bar x+n\bar x\bar y$

8\.

样本	n	$\bar x$	$\bar y$	S_x	S_y	r_{xy}
I	600	5	12	2	3	0.6
II	400	7	10	3	4	0.7

(取 $\bar x=\frac{1}{n}\sum x_i$, $S_x=\sqrt{\frac{1}{n}\sum(x_i-\bar x)^2}$)

求 I+II 的 $x.y$ 之间相关系数，且小于 1.2 的相关系数

$r_{xy}=\frac{\sum_{i=1}^n (x_i-\bar x)(y_i-\bar y)}{\sqrt{\sum_{i=1}^n (x_i-\bar x)^2}\sqrt{\sum_{i=1}^n (y_i-\bar y)^2}}=\frac{\sum x_i y_i-n\bar x\bar y}{n S_x S_y}$

$\sum x_i y_i=N r_{xy} S_x S_y+N\bar x\bar y$

63.6

因此 $\sum_{group I} x_i y_i=600(0.6\times2\times3+5\times12)=38160$

$\sum_{group II} x_i y_i=400(0.7\times3\times4+7\times10)=31360$

$\sum_{group I+II} x_i y_i=69520$

对 I+II $\bar x=0.6\times5+0.4\times7=5.8$

$\bar y=0.6\times12+0.4\times10=11.2$

$S_x=4.190465$

$S_y=3.572114$

$r_{xy(I+II)}=\frac{\sum_{I+II} x_i y_i-(n_1+n_2)\overline{X_{I+II}}\,\overline{Y_{I+II}}}{(n_1+n_2)S_{x_{I+II}}S_{y_{I+II}}}=0.30463$

$<\min(r_I, r_{II})$.

10. $Y \sim N(X\beta, \sigma^2 I_n)$ 为随机向量，$X \in \mathbb{R}^{n\times p}$ $n \geq p$

求 Y 的 PDF，并求已知 Y 时 β, σ^2 的 MLE 估计

由正态分布结论)

$$p(Y|X, \beta, \sigma) = \frac{1}{\sqrt{(2\pi)^n}\,\sigma^n} e^{-\frac{1}{2\sigma^2}(Y - X\beta)^T(Y - X\beta)}$$

$$L = \ln p = -\frac{n}{2}\ln(2\pi) - n\ln\sigma - \frac{1}{2\sigma^2}\|Y - X\beta\|_2^2$$

$$\frac{\partial L}{\partial \beta} = 0 \iff \hat{\beta} = \arg\min_{\beta} \|Y - X\beta\|_2^2, \quad \underline{X^T X \hat{\beta} = X^T y}$$

$$\frac{\partial L}{\partial \sigma} = 0 \iff -\frac{n}{\sigma} + \frac{1}{\sigma^3}\|Y - X\beta\|_2^2 \Rightarrow \underline{\hat{\sigma}^2 = \frac{1}{n}\|Y - X\hat{\beta}\|_2^2}$$

(若 $\mathrm{rank}(X) = p$，$\hat{\beta} = (X^T X)^{-1} X^T y$，$\hat{\sigma}^2 = \frac{1}{n} Y^T (I - X(X^T X)^{-1} X^T) Y$)

9.

x	60	61	62	63	64	65	66	67	68	69
y	18	63	294	424	834	854	904	1054	1054	1054

$y = \frac{L}{1 + e^{a+bx}}$ 拟合．$L = 1060$ 画出原数据，拟合曲线图形．

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

记 $y' = \ln\left(\frac{L}{y} - 1\right)$，对 (x, y') 拟合

得 $\hat{a} = 67.53721$

$\hat{b} = -1.065394$

绘图见后