

No.	数学M	理科S
1	39	44
2	63	58
3	63	60
4	43	50
5	66	60
6	71	69
7	55	60
8	48	47
9	54	53
10	43	44
11	30	34

96	57	55
97	53	60
98	65	58
99	48	40
100	53	57

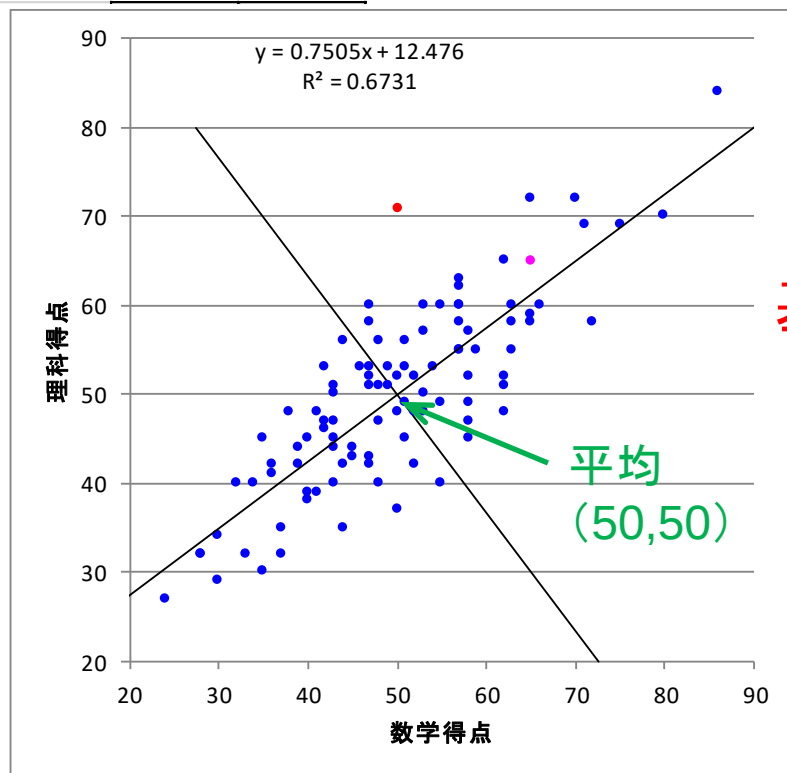
分散

$$\begin{cases} V_x = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \\ V_y = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \end{cases}$$

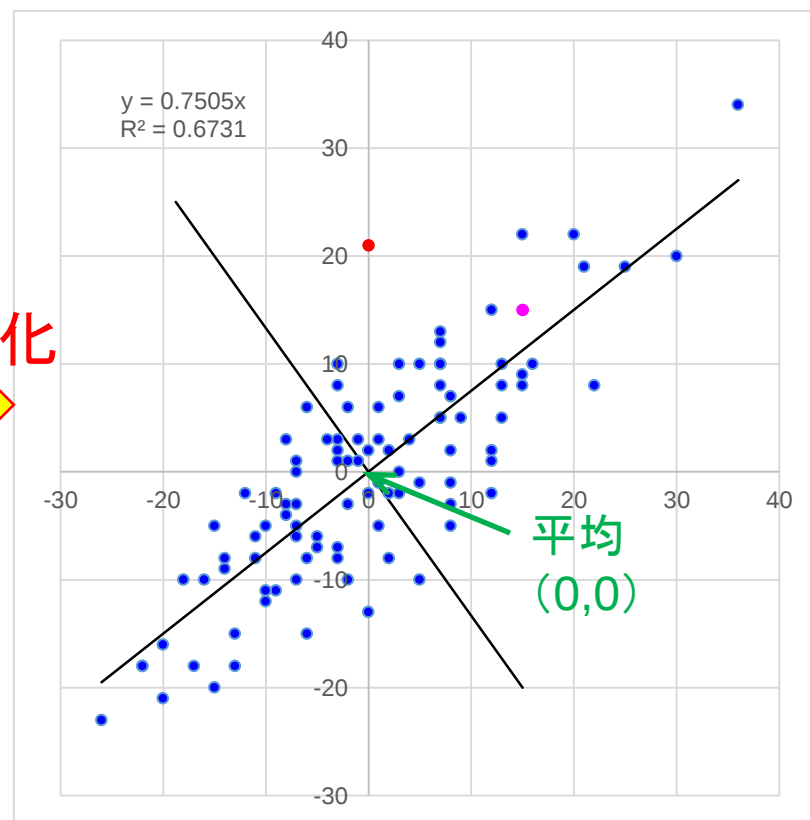
平均

共分散

$$S_{xy}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$



基準化



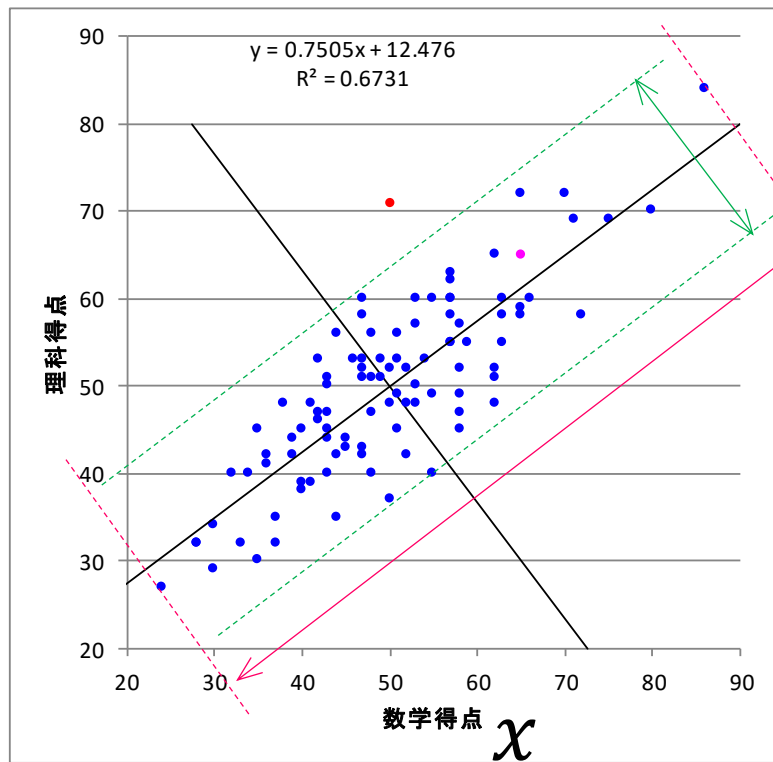
$$\text{分散} \left\{ \begin{array}{l} V_x = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \\ V_y = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \end{array} \right.$$

$$\text{共分散} \quad s_{xy}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

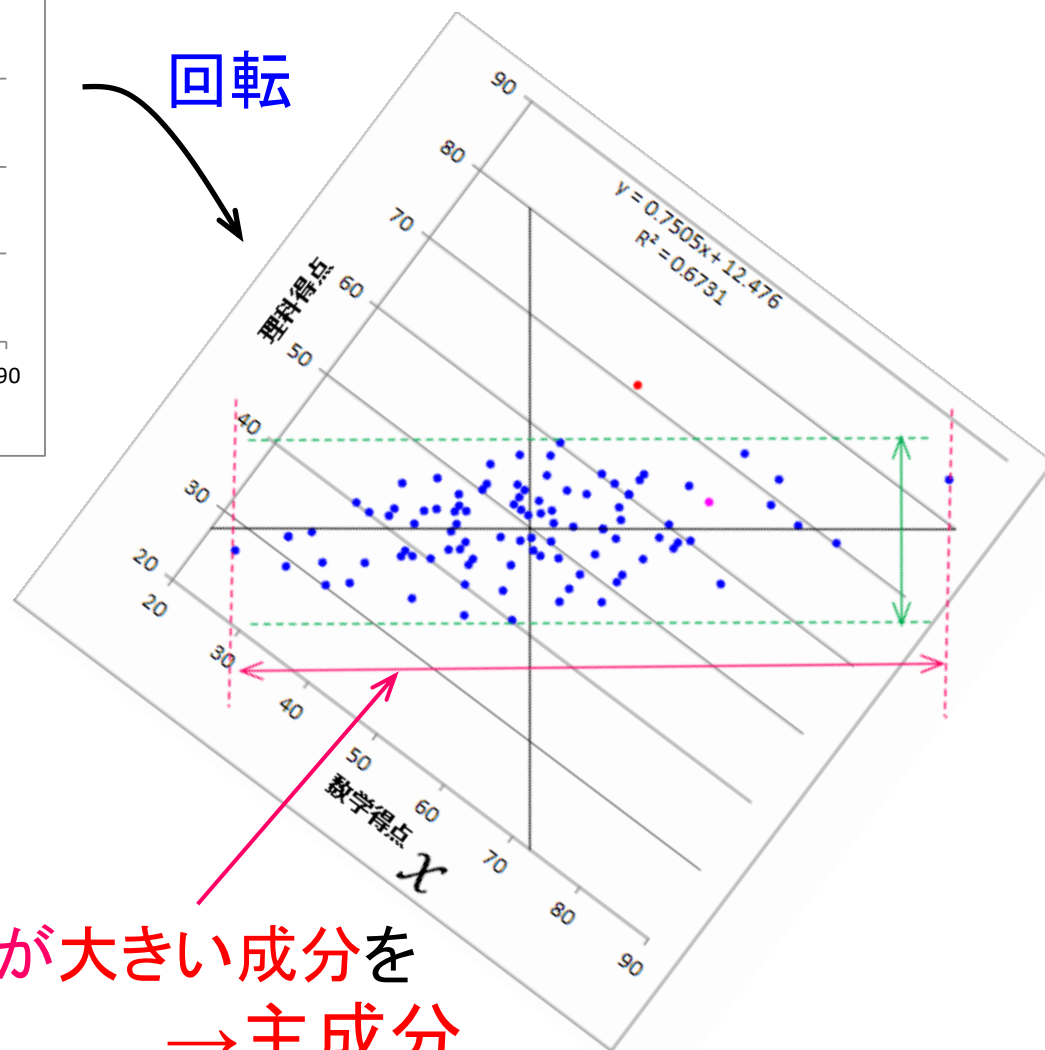
$$X = \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} \end{bmatrix}$$

$$\begin{aligned} \frac{1}{n} X^T X &= \frac{1}{n} \begin{bmatrix} x_1 - \bar{x} & \cdots & x_i - \bar{x} \\ y_1 - \bar{y} & \cdots & y_i - \bar{y} \end{bmatrix} \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} \end{bmatrix} \\ &= \frac{1}{n} \begin{bmatrix} (x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2 & (x_1 - \bar{x})(y_1 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y}) \\ (x_1 - \bar{x})(y_1 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y}) & (y_1 - \bar{y})^2 + \cdots + (y_n - \bar{y})^2 \end{bmatrix} \\ &= \begin{bmatrix} V_x & S_{xy} \\ S_{xy} & V_y \end{bmatrix} \end{aligned}$$

分散共分散行列



回転



ばらつき(分散)が大きい成分を
 x 軸にする
 →主成分

xy 平面を θ 回転させた青線の軸に
投影させた成分を z_i とおく

$$z_i = ax_i + by_i$$

$$a = \cos \theta$$

$$b = \sin \theta \quad \text{とすると} \quad a^2 + b^2 = 1$$

z_i のばらつき(分散)を V_z とすると

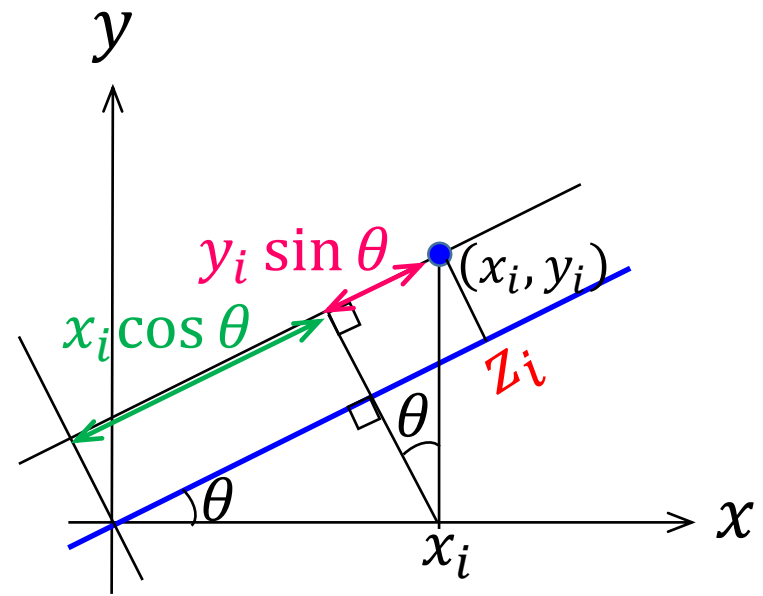
$$V_z = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2$$

$$= \frac{1}{n} \sum_{i=1}^n \{(ax_i + by_i) - (a\bar{x} + b\bar{y})\}^2$$

$$= \frac{1}{n} \sum_{i=1}^n \{a(x_i - \bar{x}) + b(y_i - \bar{y})\}^2$$

$$= \frac{1}{n} \sum_{i=1}^n \{a^2(x_i - \bar{x})^2 + 2ab(x_i - \bar{x})(y_i - \bar{y}) + b^2(y_i - \bar{y})^2\}$$

$$= a^2 V_x + 2ab S_{xy} + b^2 V_y$$



分散 V_z が大きい場合の a 、 b を算出すると、主成分が明らかになる
以下の2式よりラグランジュの未定係数法を適用する

$$\begin{cases} V_z = a^2 V_x + 2ab S_{xy} + b^2 V_y \\ a^2 + b^2 = 1 \end{cases}$$

$$F(\mathbf{a}, \mathbf{b}, \lambda) = a^2 V_x + 2ab S_{xy} + b^2 V_y - \lambda(a^2 + b^2 - 1)$$

$$\frac{\partial F(\mathbf{a}, \mathbf{b}, \lambda)}{\partial a} = 2a V_x + 2b S_{xy} - 2\lambda a = 0$$

$$\frac{\partial F(\mathbf{a}, \mathbf{b}, \lambda)}{\partial b} = 2a S_{xy} + 2b V_y - 2\lambda b = 0$$

$$\frac{\partial F(\mathbf{a}, \mathbf{b}, \lambda)}{\partial \lambda} = -a^2 - b^2 + 1 = 0$$

$$\begin{bmatrix} aV_x + bS_{xy} \\ aS_{xy} + bV_y \end{bmatrix} = \begin{bmatrix} V_x & S_{xy} \\ S_{xy} & V_y \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \lambda \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\begin{bmatrix} V_x & S_{xy} \\ S_{xy} & V_y \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} V_x - \lambda & S_{xy} \\ S_{xy} & V_y - \lambda \end{bmatrix} = (V_x - \lambda)(V_y - \lambda) - S_{xy}^2 = 0$$

$$(V_x - \lambda)(V_y - \lambda) - S_{xy}^2 = 0$$

$$\lambda^2 - (V_x + V_y)\lambda + V_x V_y - S_{xy}^2 = 0 \text{ より}$$

固有値 λ_1 及び λ_2 を算出する

$$V_x = 133.5$$

$$V_y = 111.7$$

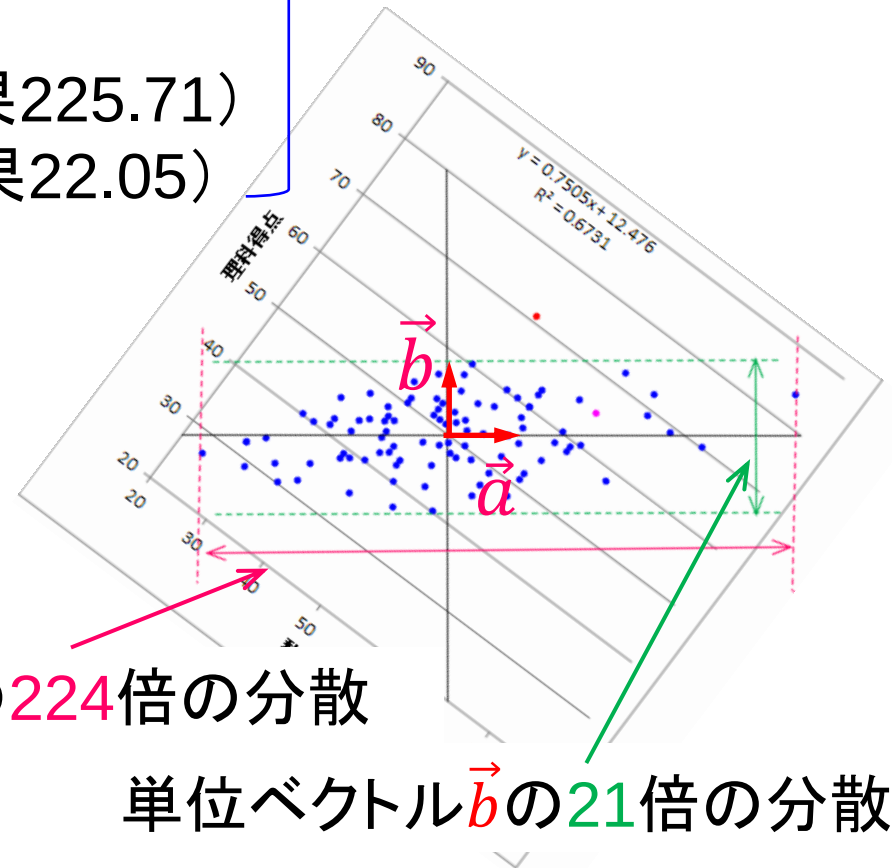
$$S_{xy} = 100.2 \text{ を代入して}$$

$$\lambda_1 = 224 \text{ (Minitabの解析結果225.71)}$$

$$\lambda_2 = 21 \text{ (Minitabの解析結果22.05)}$$

Excel参照

以上の結果より、数学の採点が第1主成分であることが明らか

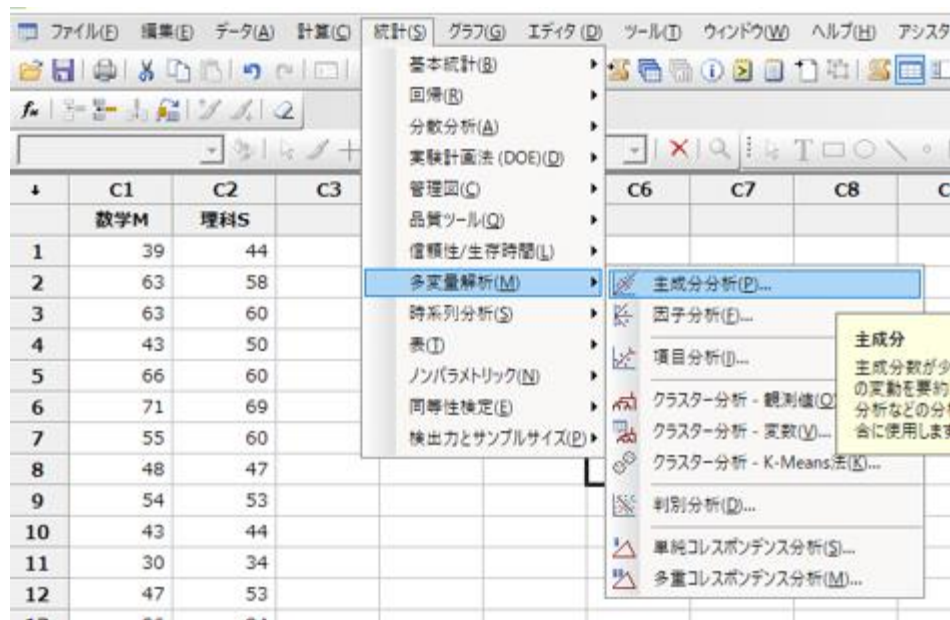


Minitabの主成分分析

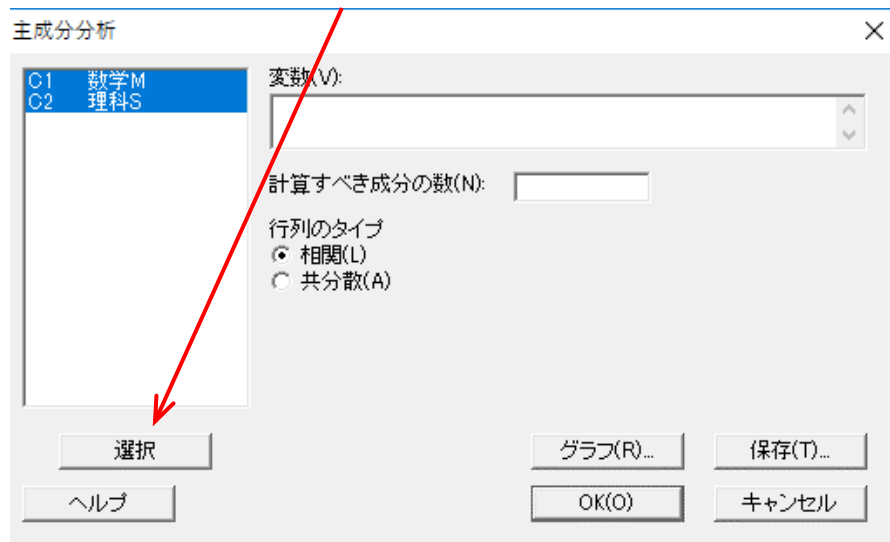
①C1に数学の採点
C2に理科の採点データを入力

	C1	C2	C3	C4
	数学M	理科S		
1	39	44		
2	63	58		
3	63	60		
4	43	50		
5	66	60		
6	71	69		
7	55	60		
8	48	47		
9	54	53		
10	43	44		
11	30	34		
12	47	53		
13	86	84		
14	51	45		
15	47	52		
16	50	52		
17	50	48		
18	37	32		
19	47	42		
20	43	45		

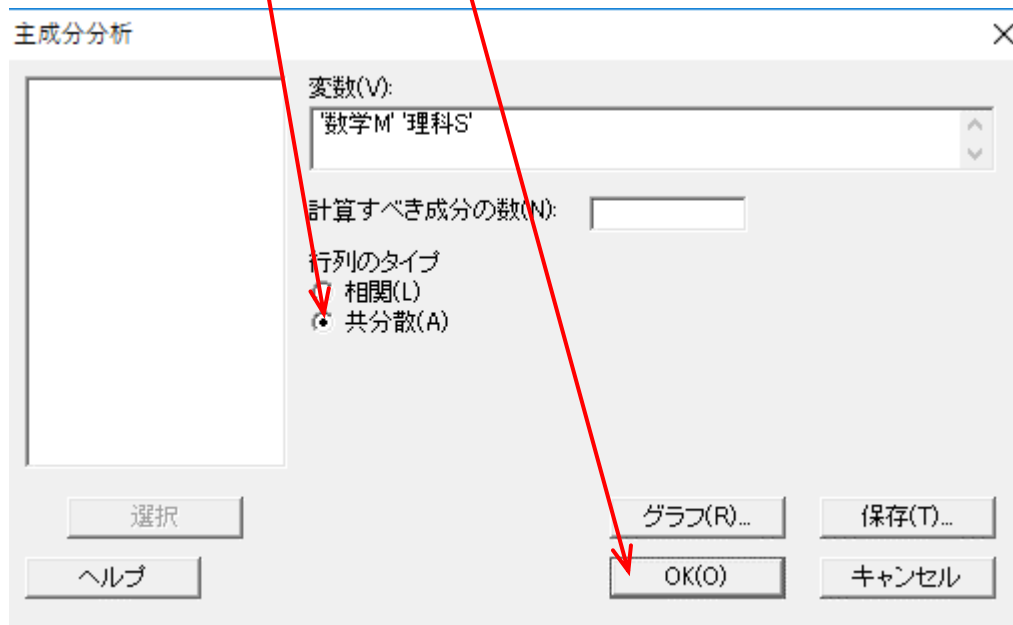
②多変量解析→主成分分析をクリック



③この2つを選択してクリック



④ 共分散選択してOKをクリック



実行結果

----- 2020/01/28 14

Minitabへようこそ、ヘルプを表示するに

主成分分析: 数学M, 理科S

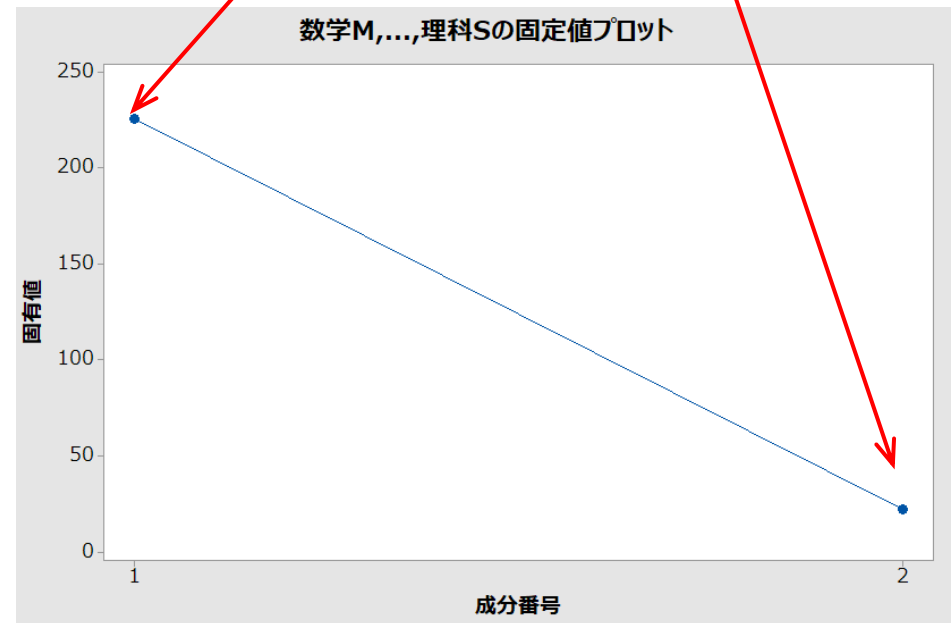
共分散行列の固有分析

固有値	225.71	22.05
比率	0.911	0.089
累積	0.911	1.000

← λ_1 、 λ_2

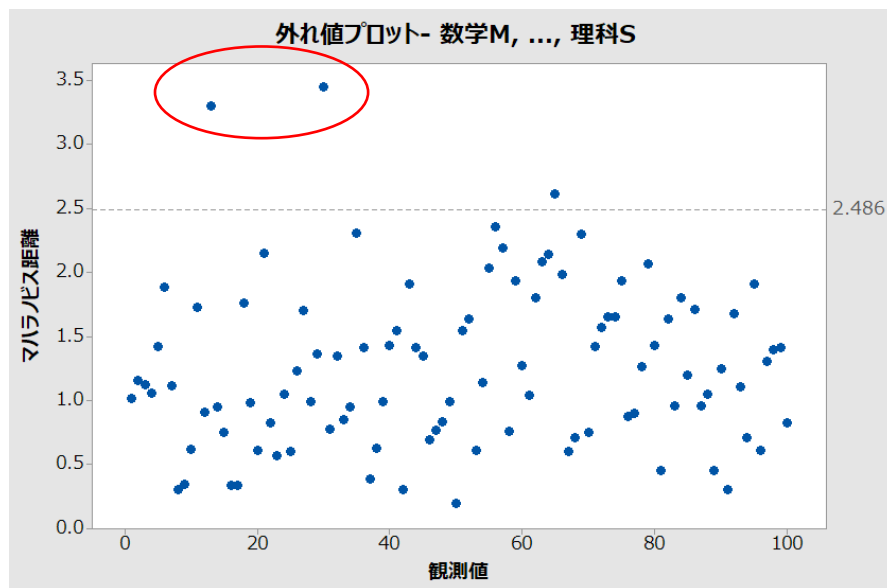
変数	PC1	PC2
数学M	0.744	-0.668
理科S	0.668	0.744

⑤ グラフを選択sh字いてOKをクリックすると右図のようなグラフ→



こんなグラフも描いてくれます

マハラノビス距離による
外れ値プロット



数学と理科のどちらの相関が
高いかを矢印の方向で示す

