

R 語言問卷分析(3)

因素分析

量表效度/信度分析

吳漢銘

國立政治大學 統計學系



<https://hmwu.idv.tw>

■ 因素分析

- 是一種(變數)維度縮減 (dimension reduction) 的技術。
- 目的是將具有多變數之資料，縮減成較少的變數，同時保持原資料所提供之大部份資訊。
- 變數縮減後，較易進行解釋或繪圖。可做後續之檢定、迴歸分析、判別分析、集群分析等等之依據。
- 因素分析結果是一個中間過程，用以**產生後階段分析所需之變數**。

本講義部份內容掃描自以下所列之上課用書，並無它用。
吳明隆, SPSS操作與應用：問卷統計分析實務 (附光碟) 五南出版社, 出版日期 / 2010/10/05(2版 5刷)。

- 項目分析完，應進行「因素分析」(共同因素分析 common factor analysis) 以求得量表之「建構效度」。
 - 「**建構效度**」：量表能測量理論概念或特質之程度。
 - 「**因素分析**」：目的在找出量表潛在的結構。
- 「效度」(validity): 能夠測到該測驗所欲測特質的程度。
 - 內在效度 (internal): 指**研究敘述**的正確性與真實性。
 - 外在效度 (external): 研究**推論**的正確性。

提高內在效度

- 在研究設計時，需
 - **理論正確解釋清楚**: 概念要具體明確，解釋要信而可徵。
 - **操作忠實減少誤差**: 概念及變項依理論建構，給予操作型的定義，設計有效度的評量工具(或測驗)。
 - **樣本合宜**: 要注意不同組別人數的相等性。重視研究情境的適當性與問卷回收率。
 - **排除無關變因**: 排除足以混淆結論之無關干擾變項。

- 操作型定義 (operational definition) ，是指將一些事物如變量、術語與客體等以某種操作或觀測的方式表示出來，與概念型定義相區別。
- 所謂「操作定義」，是定義中包含有測量方法。比如「長度」的定義包含以公里、公尺、公分等為單位，和用尺做工具來測量長度的數量；「時間」的定義包含以年、月、日、時、分、秒等為單位，和用鐘錶做工具，來測量時間的數量，所以「長度」和「時間」都是操作定義。此外，「美」和「神聖」的定義沒有包含單位和測量的方法，「人命值多少」的定義中也沒有大家共同接受的測量方法，所以「美」、「神聖」和「人命值多少」不是操作定義。

提高外在效度

- **資料收集:** 觀察具普遍性，資料來源要多元性且客觀。
- **取樣:** 有足夠代表性。
- 以**操作型定義**代表概念性定義。
- **研究情境要適切:** 最好能與未來實際應用或推論的情境類似。
- **排除無關干擾變項:** 防止實驗者效應。
- **解釋分析:** 應具普遍性，客觀性，中立性，合理性與真實性。

效度具有的性質

- 效度是指「**測驗結果**」之正確性或可靠性，而非指測驗工具本身。
- 效度並非全有或全無，只是**程度上有高低**之差別。
- 效度具**目標功能性**，未具普遍性。
 - 一高效度量表有其適用的特定群體及特殊目的。
 - 若施測不同受試者，可能導致結果不正確。
- 效度無法實際測量，僅能從現有資訊做**邏輯推論**或從實証資料做統計**考驗分析**得出。
- 測驗是否有效的歷程，在於累積証據以支持適當的，有意義及有用的**推論的過程**。

效度的分類 (1)

1. 內容效度 (content validity):

- 量表(測驗)內容(題目) 的適切性與代表性。
- 要測量的心理特質，測驗內容可否達到。
- 內容效度的檢核通常透過「雙向細目表」。
- 也常以題目分佈的合理性判斷。
- 也稱為「邏輯效度」

2. 效標關聯效度(criterion-related validity)

- 指測驗與外在效標間關係的程度。
- 常見的外在效標 (具有良好的信度與效度): 標準化之學業成就，智力測驗，人格量表，態度量表等等。
- 依時間分成兩類:
 - 「同時效度(concurrent validity)」:指測驗分數與目前的效標之間關係的程度。
 - 「預測效度(predictive validity)」:指測驗分數與將來的效標之間關係的程度。
- 效標關聯效度通常求實際測驗分數與效標間的關係，又稱「實徵性效度 (empirical validity)」

效度的分類 (2)

3. 建構效度 (construct validity):

- 實際測驗分數能解釋某一心理特質有多少。
- 測驗能夠量到理論上之建構心理特質之程度。
- 建構效度考驗步驟:
 - 根據文獻探討及前人研究結果，建立假設性理論。
 - 編製適切測驗工具。
 - 選取適當受試者進行施測。
 - 進行因素分析以了解此份測驗工具能否有效解釋所欲建構的心理特質。
 - 若所得之「共同因素」與理論架構之心理特質相近，則此測驗工具或量表具有「建構效度」。

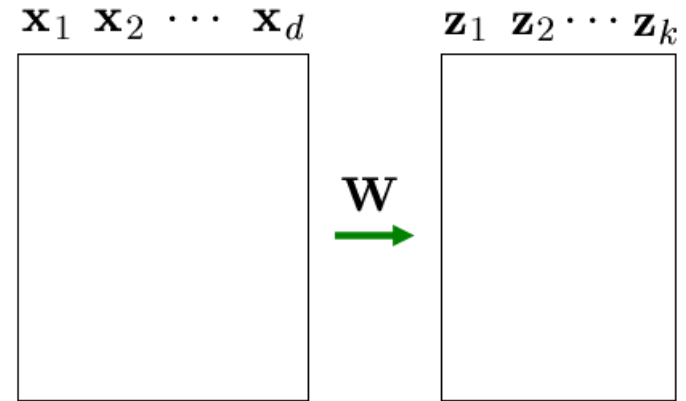
4. 專家效度

- 將編製好之量表請相關之專家檢視。
- 根據專家意見，統計分析適合之題項，再編製成預試問卷。

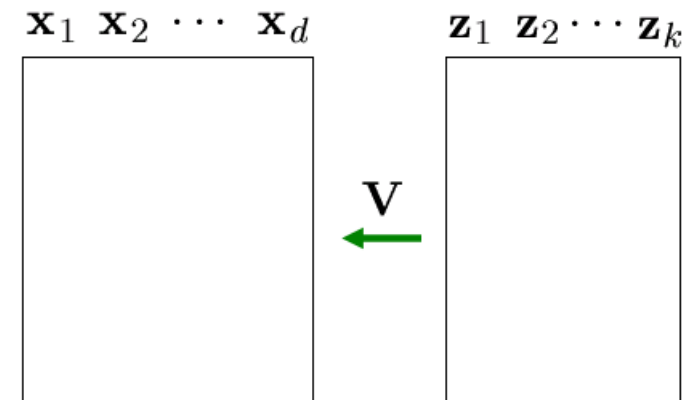
驗證性因素分析 (Confirmatory FA): 驗證理論結構，因素結構已知。

- There is a set of unobservable, **latent factors** $z_j, j = 1, \dots, k$, which when acting in combination generate x .
- **Goal:** to characterize the dependency among the observed variables by means of a smaller number of factors (groupings).
- A few factors can represent the groups of variables.
- FA always partitions the variables into factor clusters.

PCA $z = W^T(x - \mu)$



FA $x - \mu = Vz + e$





探索性因素分析 (Exploratory FA)

- FA assumes that each input dimension can be written as a weighted sum of the k factors, plus the residual term.

$$x_i - \mu_i = \sum_{j=1}^k v_{ij}z_j + \epsilon_i, \quad i = 1, \dots, d \quad \rightarrow \quad \text{Var}(x_i) = \sum_{j=1}^k v_{ij}^2 + \psi_i$$

Factor loadings



Model: $\mathbf{x} - \boldsymbol{\mu} = \mathbf{V}\mathbf{z} + \mathbf{E}$

$[d \times 1] \quad [d \times k] \quad [k \times 1] \quad [d \times 1]$

The variance explained by the common factor and unique factor

Sample

$$\mathbf{X} = \{\mathbf{x}^t\}_{t=1}^N$$

$$E[\mathbf{x}^t] = \boldsymbol{\mu}$$

$$\text{Cov}[\mathbf{x}^t] = \boldsymbol{\Sigma}$$

Common Factors

$$E[z_i] = 0, \quad i = 1, \dots, k$$

$$\text{Var}(z_i) = 1$$

$$\text{Cov}(z_i, z_j) = 0, \quad i \neq j$$

Unique Factors

$$E[\epsilon_i] = 0, \quad i = 1, \dots, k$$

$$\text{Var}(\epsilon_i) = \psi_i$$

$$\text{Cov}(\epsilon_i, \epsilon_j) = 0, \quad i \neq j$$

$$\text{Cov}(\epsilon_i, z_j) = 0, \quad \forall i, j$$

$$\mathbf{x} - \mu = \mathbf{V}\mathbf{z} + \mathbf{E}$$

$$\Sigma = \text{Cov}(\mathbf{x})$$

$$= \text{Cov}(\mathbf{V}\mathbf{z} + \mathbf{E})$$

$$= \text{Cov}(\mathbf{V}\mathbf{z}) + \text{Cov}(\mathbf{E})$$

$$= \mathbf{V}\text{Cov}(\mathbf{z})\mathbf{V}^T + \Psi$$

$$= \mathbf{V}\mathbf{V}^T + \Psi$$

$$\Psi = \text{diag}[\psi_i]_{i=1}^d$$

Example: two factors

$$\text{Cov}(x_1, x_2) = v_{11}v_{21} + v_{12}v_{22}$$

- If x_1 and x_2 have **high covariance**, then they are related through **a factor**.
- 1st factor $\Rightarrow v_{11}, v_{21}$ high
- 2nd factor $\Rightarrow v_{12}, v_{22}$ high
- In either case, the sum will be high.
- If x_1 and x_2 have **low covariance**, then they depend on **different factors** and in the products in the sum.
- One term will be high, the other will be low, the sum will be low.

Variance Component:

$$x_1 = v_{11}z_1 + v_{12}z_2 + \epsilon_1$$

$$\text{Cov}(x_1, z_2) = \text{Cov}(v_{12}z_2, z_2) = v_{12}\text{Var}(z_2) = v_{12} \quad \longrightarrow \quad \text{Cov}(\mathbf{x}, \mathbf{z}) = \mathbf{V}$$

The **loadings** represent the **correlations** of **variables** with the **factors**.



探索性因素分析 (Exploratory FA)

Model: $\mathbf{x} - \mu = \mathbf{V}\mathbf{z} + \mathbf{E}$

Variance Component: $\Sigma = \mathbf{V}\mathbf{V}^T + \Psi \longrightarrow \Psi = \text{diag}[\psi_i]_{i=1}^d$

$$\mathbf{S} = \mathbf{V}\mathbf{V}^T + \Psi$$



To find $\mathbf{V} = [d \times k]$

$$\Psi = \psi\mathbf{I} \text{ Probabilistic PCA}$$

$$\Psi = 0\mathbf{I} \text{ Conventional PCA}$$

- If there are only a few factors (\mathbf{V} has few columns), then we have a simplified structure for \mathbf{S} .
- Finding factor loadings and the specific variances.

$$\mathbf{S} = \mathbf{C}\mathbf{D}\mathbf{C}^T = \mathbf{C}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{C}^T = (\mathbf{C}\mathbf{D}^{1/2})(\mathbf{C}\mathbf{D}^{1/2})^T$$

$$\text{Var}(x_i) = \sum_{j=1}^k v_{ij}^2 + \psi_i$$



$$\psi_i = s_i^2 - \sum_{j=1}^k v_{ij}^2$$

$$\mathbf{C} = [d \times k]$$

$$\mathbf{D} = [k \times k]$$

$$\mathbf{V} = \mathbf{C}\mathbf{D}^{1/2}$$

- It can be used for knowledge extraction when we find the **loadings** and try to express the variables using fewer **factors**.

- It can also be used for dimension reduction when $k < d$.

- Find the **factor scores** z_j from \mathbf{x}_i .

$$x_i - \mu_i = \sum_{j=1}^k v_{ij} z_j + \epsilon_i, \quad i = 1, \dots, d$$

- Find the **loadings** w_{ji} such that

$$z_j = \sum_{i=1}^d w_{ji} x_i + \epsilon_i, \quad j = 1, \dots, k$$

$$\mathbf{z}^t = \mathbf{W}^T \mathbf{x}^t + \mathbf{e}, \quad \forall t = 1, \dots, N$$

NOTE: For dimension reduction, FA offers no advantage over PCA except the **interpretability** of factors allowing the identification of **common causes**, a simple explanation, and knowledge extraction.

$$\mathbf{Z} = \mathbf{XW} + \mathbf{E} \quad \rightarrow \quad \mathbf{W} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Z}$$

$$\mathbf{Z} = [N \times k]$$

$$\mathbf{W} = (N - 1)(\mathbf{X}^T \mathbf{X})^{-1} \frac{\mathbf{X}^T \mathbf{Z}}{N - 1}$$

$$\mathbf{X} = [N \times d]$$

$$\mathbf{E} = [N \times k]$$

$$\mathbf{W} = \left(\frac{\mathbf{X}^T \mathbf{X}}{N - 1} \right)^{-1} \frac{\mathbf{X}^T \mathbf{Z}}{N - 1} \quad \text{Cov}(\mathbf{x}, \mathbf{z}) = \mathbf{V}$$

$$\mathbf{W} = \mathbf{S}^{-1} \mathbf{V} \quad \rightarrow \quad \hat{\mathbf{Z}} = \mathbf{XW} = \mathbf{XS}^{-1} \mathbf{V}$$

FA: Rotation

- When V is multiplied with any orthogonal matrix ($TT^T = I$), the solution is not unique.

$$S = VV^T + \Psi \quad \longrightarrow \quad S = (VT)(VT)^T = VTT^TV^T = VIV^T = VV^T$$

- If T is an orthogonal matrix, the distance to the origin does not change.

$$z^T z = (Tx)^T (Tx) = x^T x$$

- Multiplying with an orthogonal matrix has the effect of **rotating the axes** which allows us to choose the set of axes **most interpretable**.

$$T = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \quad \text{Rotates the axes by phi}$$

- **Orthogonal rotation**: the factors are still orthogonal after the rotation.
 - ◆ **Varimax**: tries to **maximize variance** of **squared loadings** for each factor (orthogonal):
 - lines up factors with original variables.
 - improves interpretability of factors.
 - ◆ **Quartimax**: **maximizes the variance** of the **squared loadings within the variables**.
- **Oblique rotation**: the factors are allowed to become correlated.

- 因素分析可以抽取變項間的**共同因素**，以較少的再構念代表原來較複雜的資料結構。
- 因素分析假定每個指標(即題項，問卷問題)均由兩個部份組成：

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + a_{j3}F_3 + \cdots + a_{jm}F_m + U_j$$

- **共同因素 (common factor):**

- 共同因素比題項少。
- 直交轉軸法: 共同因素彼此沒有相關。
- 斜交轉軸法: 共同因素彼此有相關。

- **唯一因素 (unique factor):**

- 每個題項皆有一個唯一因素。
- 所有唯一因素間彼此沒有關係。
- 所有唯一因素與所有的共同因素也沒有關係。

- | |
|--|
| <ol style="list-style-type: none"> 1. Z_j: 第 j 個變項的標準化分數。 2. F_i: 共同因素。 3. m: 所有變項共同因素的數目。 4. U_j: 變項 Z_j 的唯一因素。 5. a_{ji}: 因素負荷量 (loading)。
第 i 個共同因素對第 j 個變項之貢獻。 |
|--|

- 假設有三個變項，抽取兩個共同因素：

$$Z_1 = a_{11}F_1 + a_{12}F_2 + U_1$$

$$Z_2 = a_{21}F_1 + a_{22}F_2 + U_2$$

$$Z_3 = a_{31}F_1 + a_{32}F_2 + U_3$$

變項	F_1	F_2	共同性 (h^2)	唯一因素 (d^2)
X_1	a_{11}	a_{12}	$a_{11}^2 + a_{12}^2$	$1 - h_1^2$
X_2	a_{21}	a_{22}	$a_{21}^2 + a_{22}^2$	$1 - h_2^2$
X_3	a_{31}	a_{32}	$a_{31}^2 + a_{32}^2$	$1 - h_3^2$
特徵值	$a_{11}^2 + a_{21}^2 + a_{31}^2$	$a_{12}^2 + a_{22}^2 + a_{32}^2$		
解釋量	$(a_{11}^2 + a_{21}^2 + a_{31}^2)/3$	$(a_{12}^2 + a_{22}^2 + a_{32}^2)/3$		

共同性與特徵值 (2)

■ 共同性 (communality):

- 是每個變項在每個共同因素之負荷量的平方總和。
- 是個別變項可以被共同因素解釋的變異量百分比。
- 是個別變項與共同因素間多元相關的平方。
- 代表所有共同因素對第j個變項變異量所能解釋的部份。
- 共同性愈大表示此變項與共同因素間的關係程度愈高。

■ 唯一因素:

- 大小就是1-共同性。
- 註: PCA中, 沒有唯一因素, 有多少變項, 就有多少成份。共同性等於1。

■ 特徵值 (eigenvalue):

- 是每個變項在某一共同因素之因素負荷量的平方總和。
- 因素分析的共同因素抽取中, 特徵值最大的共同因素會先被抽取。
- 以最少的共同因素, 能對總變異量做最大的解釋。
- 特徵值表示成份的變異量。

因素分析的應用

- 顯示變項間因素分析的樣態 (pattern)。
- 偵測變項間之群組 (clusters)。
- 減少大量變項數目，使之成為一組包含較少變項的統計自變項 (稱為因素)。

- 因素分析後因素數值後續應用處理方式
 - 迴歸使用: 使用標準化數值
 - 多變量分析使用: 使用各因素構面包含項目原始數值的平均值

因素分析的步驟 (1)

1. 計算變項間「相關矩陣」或「共變數矩陣」

- 若一變項與其它變項相關很低，可考慮剔除。
- 需同時考量變項的共同性與因素負荷量。

2. 估計因素負荷量

- 主成份分析法(principal components analysis)
- 主軸法
- 一般化最小平方法
- 未加權最小平方法
- 最大概似法
- Alpha因素抽取法
- 映象抽取法

3. 決定轉軸方法 (rotation)

- 目的在於改變題項在各因素之負荷量大小。
- 使因素負荷量易於解釋。
- 轉軸後，大部份的題項在每個共同因素中有一個差異較大的因素負荷量。
- 轉軸前後，共同性不變。

4. 決定因素與命名

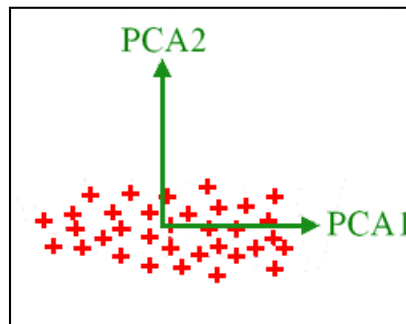
- 轉軸後，要決定因素數目，選取較少因素層面，獲得較大的解釋量。
- 因素分析程序(以兩個共同因素為例)
 - 第一次儲存之因素分數的變項名稱: 「FAC1_1」 「FAC1_2」。
 - 第二次儲存之因素分數的變項名稱: 「FAC2_1」 「FAC2_2」。

主成份分析法

- 以線性方程式將所有的變項合併，計算所有變項共同解釋的變異量。此線性組合稱為「主要成份」。
- 第一個線性組合(第一個主成份)所解釋的變異量最大。再從所剩餘的變異量中，抽出第二個主成份。
- **PCA** is a method that reduces data dimensionality by finding the new variables (major axes, principal components).
- Amongst all possible projections, PCA finds the projections so that the **maximum amount of information**, measured in terms of **variability**, is retained in the smallest number of dimensions.

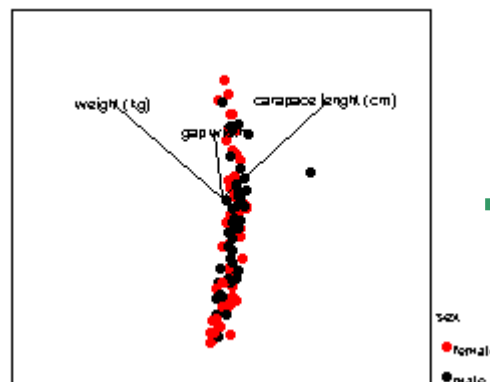
$$PCA_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p$$

$$PCA_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p$$



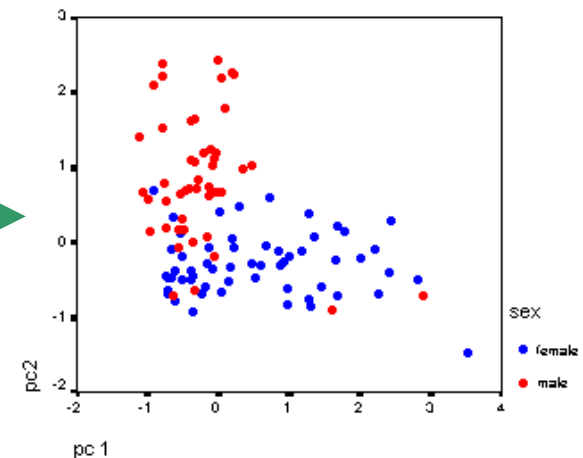
$$PCA_1 = a_1X + b_1Y$$

$$PCA_2 = a_2X + b_2Y$$



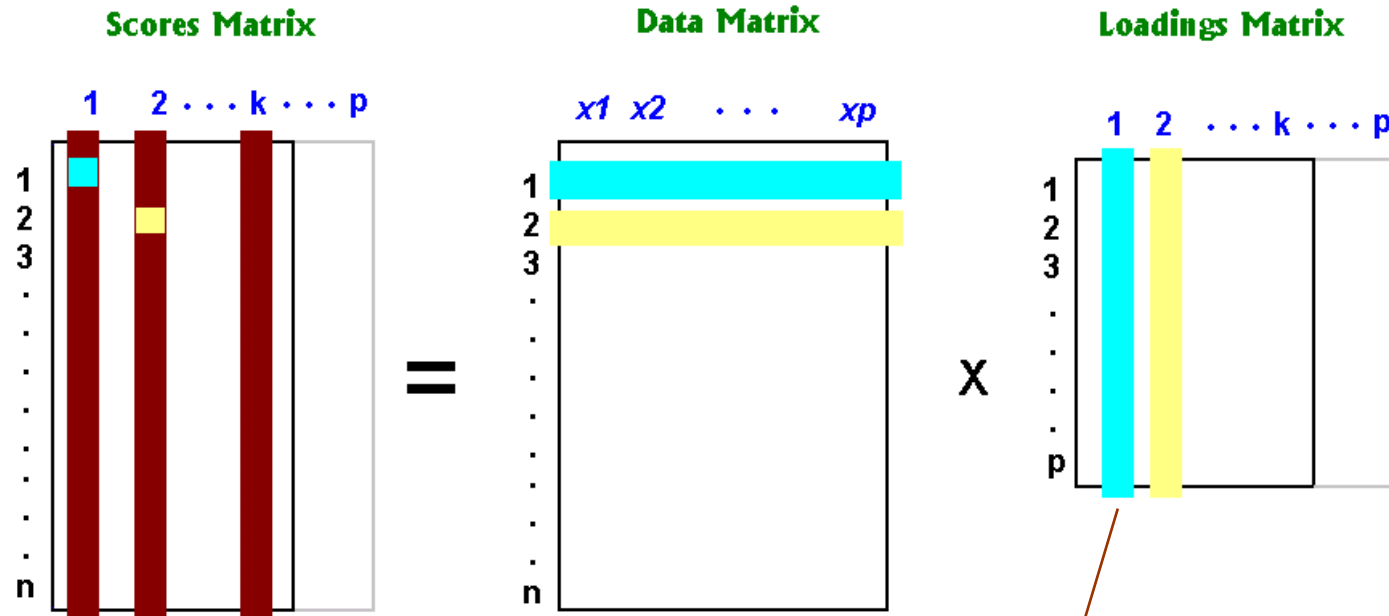
$$PCA_1 = a_1X + b_1Y + c_1Z$$

$$PCA_2 = a_2X + b_2Y + c_2Z$$



PCA: Loadings and Scores

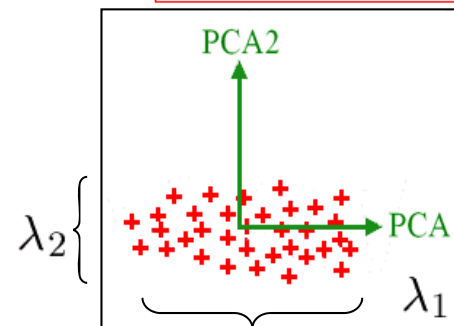
$$Z = XW$$



The i th principal component of X is Xw_i , where w_i is the i th normalized eigenvector of Σ_x corresponding to the i th largest eigenvalues.

Eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$

$$\text{proportion} = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^p \lambda_i}$$



因素負荷量

- 樣本大小與因素負荷量選取的標準:
- 因素負荷量、解釋變異百分比及題項選取準則:

樣本大小	因素負荷量選取標準值
50	0.750
60	0.700
70	0.650
85	0.600
100	0.550
120	0.500
150	0.450
200	0.400
250	0.350
350	0.300

因素負荷量	因素負荷量 ² (解釋變異量)	題項變數狀況
.71	50%	甚為理想 (excellent)
.63	40%	非常好 (very good)
.55	30%	好 (good)
.45	20%	普通 (fair)
.32	10%	不好 (poor)
<.32	<10%	捨棄

- 結論: 在因素分析中，因素負荷量最好在0.4以上，此時共同因素可以解釋題項變異的百分比為16%。

轉軸法

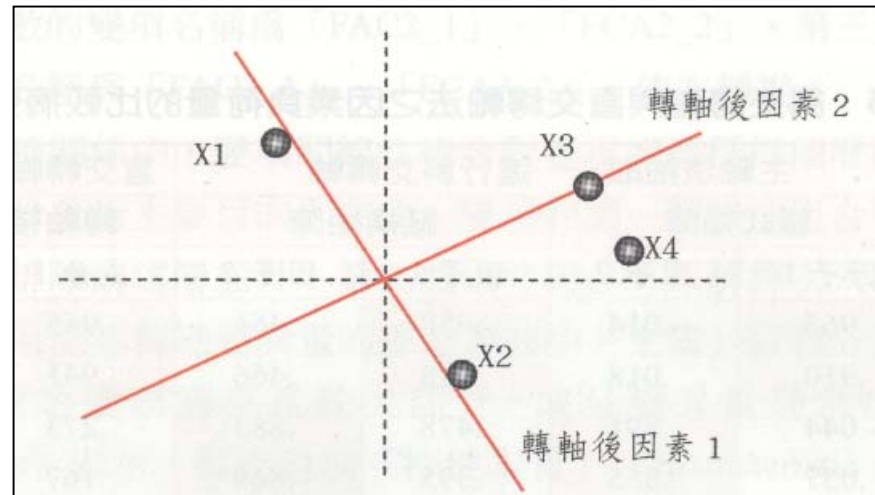
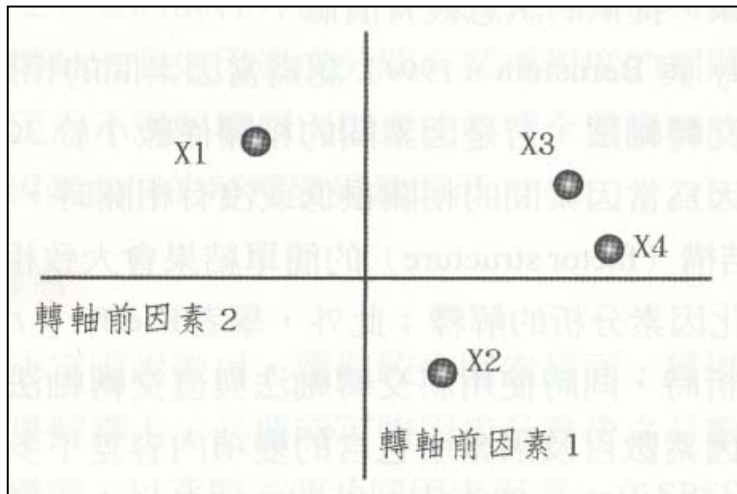
■ 直交轉軸法:

- 因素(成份)與因素(成份)沒有相關。
- 方法: **最大變異法 (varmax)**，四次方最大值法 (Quartimax)，相等最大值法 (Equamax)
- 優點: 因素間的資訊不會重疊。
- 缺點: 不符合實際生活情境。
- 最大變異法 (varmax)
 - 目的: 簡化因素矩陣的直行(共同因素)。
 - 同時每一直欄的因素負荷量平方的變異量能夠最大。

■ 斜交轉軸法:

- 因素(成份)與因素(成份)有相關。
- 方法: 直接斜交轉軸法 (Direct Oblimin), 最優轉軸法 (Promax)
- 結果產生三個矩陣: 因素結構矩陣，因素樣式矩陣，因素相關矩陣。
- 優點: 較能反應實際現象，提供有價值的訊息。
- 缺點: 解釋不易。

- 例子: 變項X1, X2可歸於共同因素一。變項X3, X4可歸於共同因素二。



	未轉軸前成分矩陣			轉軸後的成分矩陣		
	成分 1	成分 2	共同性	成分 1	成分 2	共同性
X2	.842	-.477	0.936	.941	.225	.0936
X1	.841	-.484	0.942	.945	.219	0.942
X3	.811	.469	0.878	.273	.896	0.878
X4	.757	.566	0.893	.167	.930	0.893
特徵值	2.647	1.002		1.881	1.766	

- 當因素的相關係數在0.3以上時，採用斜交轉軸法。反之，採用直交轉軸法。
- 同時比較「直交轉軸法」與「斜交轉軸法」，若因素個數及因素包含之變項內容差異不大，則採用「直交轉軸法」。若差異大，則採用「斜交轉軸法」。

斜交轉軸與直交轉軸法之因素負荷量的比較摘要表

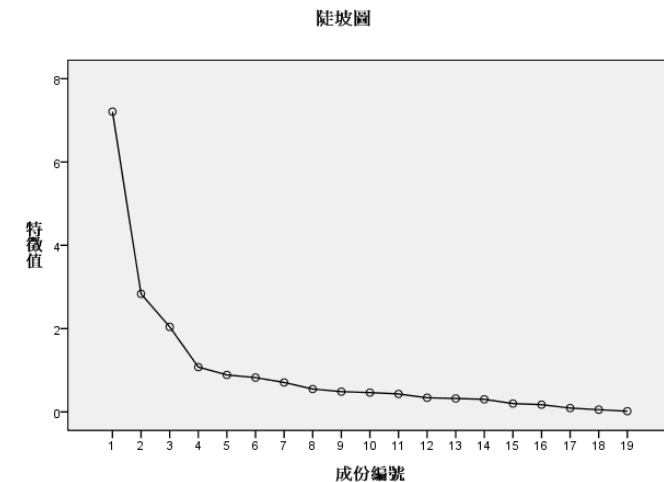
	主軸法抽取——進行斜交轉軸				直交轉軸最大變異法	
	樣式矩陣		結構矩陣		轉軸後成分矩陣	
	因子 1	因子 2	因子 1	因子 2	成分 1	成分 2
X1	.965	-.014	.958	.461	.945	.219
X2	.910	.018	.918	.466	.941	.225
X4	-.044	.891	.478	.883	.273	.896
X3	.057	.855	.395	.869	.167	.930
特徵值			2.145	1.965	1.881	1.766

1. 特徵值大於1 (Kaiser, 1960):

- 若題項數介於10到40，此方法為可靠。
- 若題項數大於40或共同性低於0.4，此方法會造成高估因素數目。
- 若題項數介於10到15或適中(20-30)，且共同性大於0.7以上，此方法最可正確。

2. 陡坡圖考驗法 (scree plot test):

- 因素變異量為縱軸，因素數目為橫軸。
- 若圖形呈現由斜坡轉為平坦，則平坦以後之共同因素可刪除。
- 當樣本大於250人，變項的共同性大於等於0.6，且因素數與變項數比小於0.3，則使用「特徵值大於1」與「陡坡圖考驗法」，才能產生精確的因素數目。



3. 變異數百分比決定法:

- 自然科學: 所選取的共同因素的累積解釋變異量至少要達95%以上。
- 社會科學: 達60%以上，表示共同因素是可靠的。

4. 事先決定法則:

- 所參考的文獻有很明確的因素構念。

5. 從相關矩陣中篩選:

- 題項間若有顯著的相關，則因素分析較易構成有意義的層面。

相關		X1	X2	X3	X4
	X1	1.000			
	X2	.879*	1.000		
	X3	.462*	.436*	1.000	
	X4	.356*	.388*	.765*	1.000

共同因素個數之選取 (3)

6. 從反映像矩陣中的MSA (measure of sampling adequacy) 篩選:

- MSA: 若某變項MSA值愈接近1 (0)，則愈(不) 適合投入因素分析中。
- $MSA > 0.8$: 題項與其它變項有共同因素存在。
- $MSA < 0.5$: 此題項不適合因素分析，可考慮刪除。

		X1	X2	X3	X4
反映像相關	X1	.582(a)	-.853	-.227	.142
	X2	-.853	.594(a)	.074	-.169
	X3	-.227	.074	.629(a)	-.725
	X4	.142	-.169	-.725	.600(a)

a 取樣適切性量數 (MSA)。

$$MSA_j = \frac{\sum_{k \neq j} r_{jk}^2}{\sum_{k \neq j} r_{jk}^2 + \sum_{k \neq j} p_{jk}^2}$$

$$KMO = \frac{\sum_{j \neq k} \sum_{j \neq k} r_{jk}^2}{\sum_{j \neq k} \sum_{j \neq k} r_{jk}^2 + \sum_{j \neq k} \sum_{j \neq k} p_{jk}^2}$$

Here r_{jk} is the correlation between the variable in question and another, and p_{jk} is the partial correlation.

反映像矩陣，即變量之間的相關矩陣的逆矩陣

Let $S^2 = (\text{diag}(R^{-1}))^{-1}$ and $Q = SR^{-1}S$. Then Q is said to be the anti-image intercorrelation matrix. Let $\text{sum}r^2 = \sum R^2$ and $\text{sum}q^2 = \sum Q^2$ for all off-diagonal elements of R and Q , then $MSA = \frac{\text{sum}r^2}{\text{sum}r^2 + \text{sum}q^2}$.

Although originally MSA was $1 - \frac{\text{sum}q^2}{\text{sum}r^2}$ (Kaiser, 1970), this was modified in Kaiser and Rice, (1974) to be $MSA = \frac{\text{sum}r^2}{\text{sum}r^2 + \text{sum}q^2}$. This is the formula used by Dziuban and Shirkey (1974) and by SPSS.

7. 由預試樣本大小來判斷:

- 受試樣本要比量表項數還多。比例5:1。
- 量表項數，不是問卷總題數，是包含項數最多的一份量表。(一份問卷可能包含多種量表)
- 建立精確的效度，樣本數最好在150人以上。
- 每個題項所需樣本最少要有5位。

8. KMO值選取法:

- Kaiser-Meyer-Olkin measure of sampling adequacy
- 依據變項間的淨(偏)相關而得 (partial correlations)
- 若變項間具有相關，則簡單相關係數會很高，但淨相關係數會較小(接近0)。
- 若兩變項的淨相關愈小，表示變項間愈有共同因素。

KMO統計量值	判別說明	因素分析適切性
.90 以上	極適合進行因素分析 (marvelous)	極佳的 (Perfect)
.80 以上	適合進行因素分析 (meritorious)	良好的 (Meritorious)
.70 以上	尚可進行因素分析 (middling)	適中的 (Middling)
.60 以上	勉強可進行因素分析 (mediocre)	普通的 (Mediocre)
.50 以上	不適合進行因素分析 (miserable)	欠佳的 (Miserable)
.50 以下	非常不適合進行因素分析 (unacceptable)	無法接受的 (Unacceptable)

- 共同因素所包含題項過於紛歧，無法命名。
 - 經多次探討與多次因素分析，逐一刪除不適切題項。
- 依特徵值大於1選取因素，與文獻理論相差過大。
 - 限制因素抽取數，以符合原先編製的架構。
- 原則: 共同因素所包含的題項同性質要高，並要能命名。



因素分析實例操作： 學校知識管理量表

- 研究:「國民中學學校知識管理與學校效能關係研究」
- 量表:「學校知識管理量表」
- 題數(原本): 20題。第11題為反向題。
- 因素分析:
 - 為探究量表的信度與題項的適切性，今隨機抽取200位國中教師為預試對象，經「項目分析」後，刪除第20題，保留19題。
 - 求此19題之建構效度。
- 注意:
 - 量表依三大面向編製:「知識創新」「知識分享」「知識獲得」。
 - 第11題反向題，需在進行「項目分析」時已反向計分。
 - 題項編號: c1,...,c19

題項	完全不符合	多數不符合	半數不符合	多數符合	完全符合
01.本校常鼓勵教師創新教學或工作創新。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02.本校教師會積極尋求班級經營上的創新。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
03.教師會積極的在其負責的行政工作上創新展現。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
04.本校教師會應用研習心得於教育品質的提升。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
05.本校會激勵教師以創新理念提升學生學習成效。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
06.本校鼓勵教師以創新有效方法激勵學生學習動機。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
07.校長會積極鼓勵同仁，分享研習吸取的新知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
08.本校教師會將班級經營的有效策略，與其他教師分享。	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
09.本校教師會在相關會議中提供意見供其他教師分享。....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.本校行政事務處理流程有完整紀錄，以供同仁分享參考。	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.本校教師很少於教學研討會上，分享其教學經驗。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12.本校同仁會於朝會上分享其研習的心得與知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13.本校教師會於同仁會議中分享其處理學生問題的策略。	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14.學校鼓勵同仁參訪標竿學校以獲取教學及行政知能。....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15.學校會鼓勵教師透過教學觀摩，以獲取專業知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.學校積極鼓勵教師參與研習活動，以獲取專業知能。....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17.學校鼓勵教師透過教師社群活動，以獲取專業知能。....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18.學校鼓勵教師透過數位資料來獲取新知識。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19.學校會影印相關教育新知給教師，以增師知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



範例一: (第一次)因素分析

33/56

因素抽取方法: minimizes the entire residual matrix using an OLS procedure + 最大變異直交轉軸法

```
fa(r, nfactors=1, rotate="oblimin", scores="regression", fm="minres", ...)
```

fm: Factoring method **fm = "minres"** will do a minimum residual as will `fm="uls"`. Both of these use a first derivative. `fm="ols"` differs very slightly from "minres" in that it minimizes the entire residual matrix using an OLS procedure but uses the empirical first derivative. This will be slower. `fm="wls"` will do a weighted least squares (WLS) solution, `fm="gls"` does a generalized weighted least squares (GLS), **fm = "pa"** will do the principal factor solution, `fm="ml"` will do a maximum likelihood factor analysis. `fm="minchi"` will minimize the sample size weighted chi square when treating pairwise correlations with different number of subjects per pair. `fm="minrank"` will do a minimum rank factor analysis. "old.min" will do minimal residual the way it was done prior to April, 2017 (see discussion below). `fm="alpha"` will do alpha factor analysis as described in Kaiser and Coffey (1965)

rotate: "none", **"varimax"**, "quartimax", "bentlerT", "equamax", "varimin", "geominT" and "bifactor" are orthogonal rotations. "Promax", "promax", "oblimin", "simplimax", "bentlerQ", "geominQ" and "biquartimin" and "cluster" are possible oblique transformations of the solution. The default is to do a oblimin transformation, although versions prior to 2009 defaulted to varimax. SPSS seems to do a Kaiser normalization before doing Promax, this is done here by the call to "promax" which does the normalization before calling Promax in GPArotation.

```
> Knowledge_Management <- read.csv("data/學校知識管理_1.csv")
> head(Knowledge_Management)
  c1 c2 c3 c4 c5 c6 c7 c8 c9 c10 c11 c12 c13 c14 c15 c16 c17 c18 c19
1  4  3  3  3  3  4  4  4  4  5  5  2  5  4  4  4  3  5  3
2  4  4  4  4  4  5  4  3  3  3  4  2  3  4  4  4  3  5  3
...
> dim(Knowledge_Management)
[1] 200 19
>
> library(psych)
> efa <- fa(Knowledge_Management, nfactors = 4, rotate = "varimax",
+          scores = "regression")
```

範例一：(第一次)因素分析

```

> efa
Factor Analysis using method = minres
Call: fa(r = Knowledge_Management, nfactors = 4, rotate = "varimax",
  scores = "regression")
Standardized loadings (pattern matrix) based upon correlation matrix
      MR1  MR3  MR2  MR4  h2  u2 com
c1  0.12  0.04  0.61 -0.09  0.402  0.598  1.1
c2  0.08  0.08  0.78 -0.04  0.628  0.372  1.0
...
c19 0.22  0.89  0.05  0.00  0.852  0.148  1.1

      MR1  MR3  MR2  MR4
SS loadings      4.26  3.57  3.07  0.88
Proportion Var   0.22  0.19  0.16  0.05
Cumulative Var   0.22  0.41  0.57  0.62
Proportion Explained 0.36  0.30  0.26  0.07
Cumulative Proportion 0.36  0.66  0.93  1.00

Mean item complexity = 1.7
Test of the hypothesis that 4 factors are sufficient.

df null model = 171 with the objective function = 16.05 with Chi Square = 3079.15
df of the model are 101 and the objective function was 2.3
...
Measures of factor score adequacy

      MR1  MR3  MR2  MR4
Correlation of (regression) scores with factors 0.97  0.99  0.96  0.82
Multiple R square of scores with factors      0.94  0.97  0.93  0.67
Minimum correlation of possible factor scores   0.88  0.95  0.85  0.34

```

- 共同性愈大表示此變項與共同因素間的關係程度愈高。
- 共同性愈小，表示此變項不適合投入主成份分析中。
- 共同性越接近1代表在項目分析中的效度指標越好，可以作為某一題目保留或是修改的標準之一。
- 考慮刪除共同性<0.2之題項。

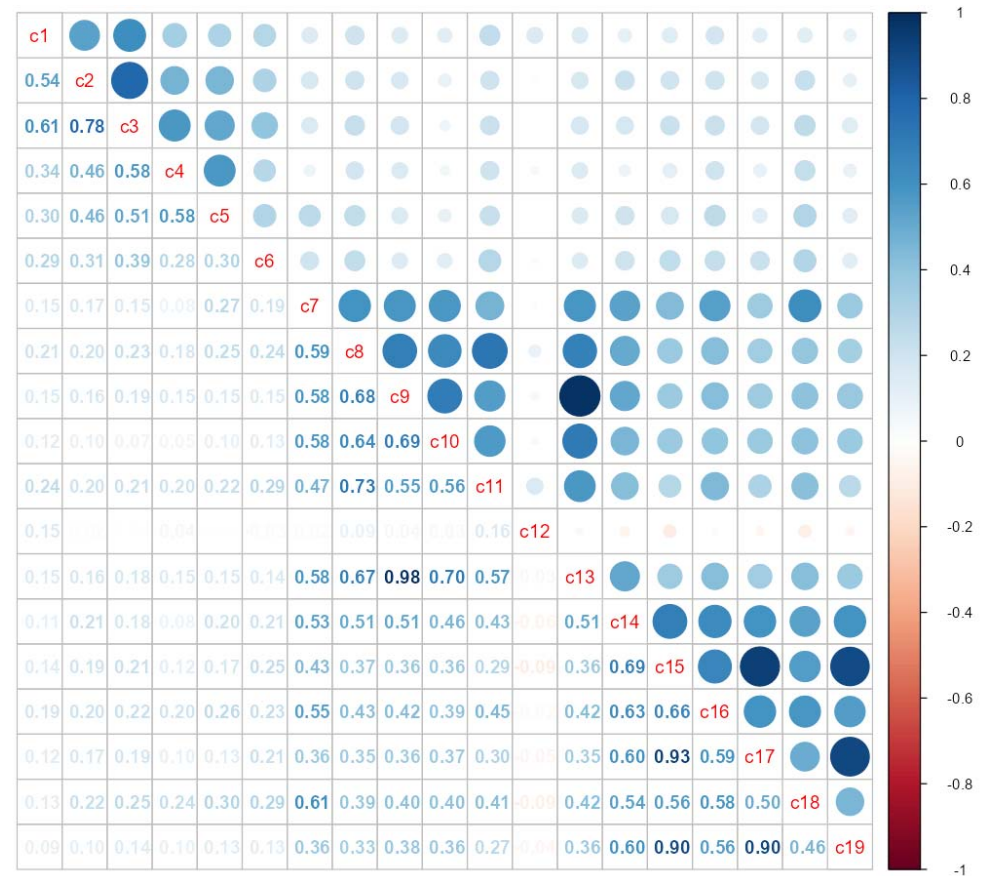
h2: efa\$communality

u2: efa\$uniquenesses

- 若有二個變項完全線性重合，則行列式為0，即無法進行因素分析。
- 如果某變項與其它變項相關係數很低，可刪除此變項。
- 變項間相關很低: 很難抽出共同因素。
- 變項間相關很高: 可能只抽出一個共同因素。
- 最好呈現某些高度相關，某些低度相關。
 - 群內高度相關: c1~c6, c7~c13 (c12除外), c14~c19。
 - 可能有共同因素存在。

```
library(corrplot)
#corrplot(cor(Knowledge_Management))
corrplot.mixed(cor(Knowledge_Management))
```

```
> cor_KM <- cor(Knowledge_Management)
> cor_ind <- which(cor_KM > 0.9, arr.ind = T)
> cor_ind[cor_ind[,1] != cor_ind[,2], ]
      row col
c13  13   9
c9   9   13
c17  17  15
c15  15  17
c19  19  17
c17  17  19
```



- KMO(值介於0~1)愈大，表示變項間共同因素愈多。
- KMO=0.855>0.8為「良好」的標準，表示變項間具有共同因素存在，適合進行因素分析。
- Bartlett 球形檢定 (Bartlett's test for sphericity):
 - H0: 變項間的淨相關不是單元矩陣。
 - H0: the correlation matrix is an identity matrix. An identity correlation matrix means your variables are unrelated and not ideal for factor analysis
 - 淨相關愈低，愈適合進行因素分析。
 - P=0.000<0.05，表示接受淨相關是單元矩陣的假設。
 - 代表母群體的相關矩陣間有共同因素存在，愈適合進行因素分析。

```
> KMO(Knowledge_Management)
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = Knowledge_Management)
Overall MSA = 0.86
MSA for each item =
  c1  c2  c3  c4  c5  c6  c7  c8  c9  c10 c11 c12 c13
0.86 0.81 0.78 0.80 0.87 0.89 0.89 0.88 0.77 0.96 0.84 0.46 0.78
  c14 c15 c16 c17 c18 c19
0.93 0.83 0.94 0.84 0.91 0.90
```

* 0.00 to 0.49 unacceptable
 * 0.50 to 0.59 miserable
 * 0.60 to 0.69 mediocre
 * 0.70 to 0.79 middling
 * 0.80 to 0.89 meritorious
 * 0.90 to 1.00 marvelous

```
> cortest.bartlett(Knowledge_Management)
R was not square, finding R from data
$chisq
[1] 3079.151
$p.value
[1] 0
$df
[1] 171
```

KMO統計量值	判別說明	因素分析適切性
.90 以上	極適合進行因素分析 (marvelous)	極佳的 (Perfect)
.80 以上	適合進行因素分析 (meritorious)	良好的 (Meritorious)
.70 以上	尚可進行因素分析 (middling)	適中的 (Middling)
.60 以上	勉強可進行因素分析 (mediocre)	普通的 (Mediocre)
.50 以上	不適合進行因素分析 (miserable)	欠佳的 (Miserable)
.50 以下	非常不適合進行因素分析 (unacceptable)	無法接受的 (Unacceptable)



Check suitability of data for Factor Analysis (FA) with Bartlett's Test of Sphericity and KMO

37/56

```
> library(performance)
> check_factorstructure(Knowledge_Management)
# Is the data suitable for Factor Analysis?
```

- Sphericity: Bartlett's test of sphericity suggests that there is sufficient significant correlation in the data for factor analysis (Chisq(171) = 3079.15, $p < .001$).

- KMO: The Kaiser, Meyer, Olkin (KMO) overall measure of sampling adequacy suggests that data seems appropriate for factor analysis (KMO = 0.86). The individual KMO scores are: c1 (0.86), c2 (0.81), c3 (0.78), c4 (0.80), c5 (0.87), c6 (0.89), c7 (0.89), c8 (0.88), c9 (0.77), c10 (0.96), c11 (0.84), c12 (0.46*), c13 (0.78), c14 (0.93), c15 (0.83), c16 (0.94), c17 (0.84), c18 (0.91), c19 (0.90).

```
> check_kmo(Knowledge_Management)
# KMO Measure of Sampling Adequacy
```

The Kaiser, Meyer, Olkin (KMO) overall measure of sampling adequacy suggests that data seems appropriate for factor analysis (KMO = 0.86). The individual KMO scores are: c1 (0.86), c2 (0.81), c3 (0.78), c4 (0.80), c5 (0.87), c6 (0.89), c7 (0.89), c8 (0.88), c9 (0.77), c10 (0.96), c11 (0.84), c12 (0.46*), c13 (0.78), c14 (0.93), c15 (0.83), c16 (0.94), c17 (0.84), c18 (0.91), c19 (0.90).

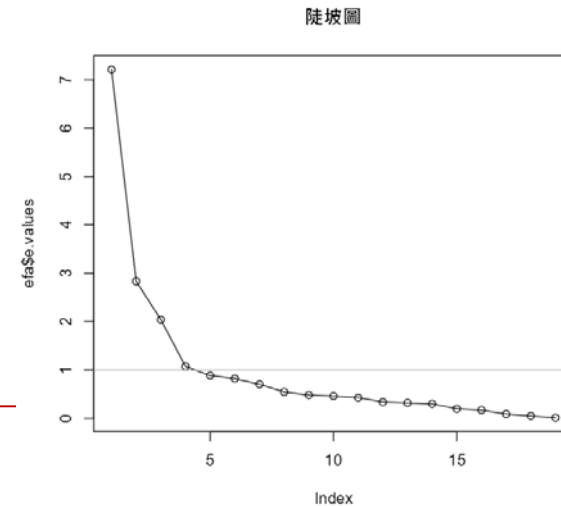
```
> check_sphericity_bartlett(Knowledge_Management)
# Test of Sphericity
```

Bartlett's test of sphericity suggests that there is sufficient significant correlation in the data for factor analysis (Chisq(171) = 3079.15, $p < .001$).

```
> efa$e.values
[1] 7.20790287 2.83361818 2.04134574 1.07528180 0.88722262 0.82362755
[7] 0.70689940 0.54845197 0.48568492 0.46372832 0.43061301 0.33915338
[13] 0.32139278 0.30155029 0.20033434 0.17249443 0.09176201 0.05331739
[19] 0.01561900
> plot(efa$e.values, type = "o", main = "陡坡圖")
> abline(h = 1, col = "grey")
```

```
> var_explained <- data.frame(特徵值 = efa$e.values,
+                             變異數的百分比 = 100 * efa$e.values/sum(efa$e.values),
+                             累積的百分比 = 100 * cumsum(efa$e.values/sum(efa$e.values)))
>
> rownames(var_explained) <- colnames(Knowledge_Management)
> round(var_explained, 3)
```

	特徵值	變異數的百分比	累積的百分比
c1	7.208	37.936	37.936
c2	2.834	14.914	52.850
c3	2.041	10.744	63.594
c4	1.075	5.659	69.253
...			
c16	0.172	0.908	99.154
c17	0.092	0.483	99.637
c18	0.053	0.281	99.918
c19	0.016	0.082	100.000



- 共同因素4，僅包含c12，可考慮刪除。
- 共同因素2中的c14, c18題項與共同因素1有密切相關。
- 表示使用「最大變異轉軸法」，未符合簡單結構要求。

```
> efa$loadings
```

Loadings:

	MR1	MR3	MR2	MR4
c1	0.124		0.614	
c2			0.783	
c3			0.937	
c4			0.642	
c5			0.604	0.263
c6	0.122	0.124	0.411	0.176
c7	0.612	0.247	0.117	0.424
c8	0.760	0.160	0.193	0.152
c9	0.911	0.192		
c10	0.749	0.212		
c11	0.650	0.122	0.220	0.194
c12		-0.101		
c13	0.916	0.181		
c14	0.451	0.557	0.118	0.262
c15	0.200	0.947	0.134	0.149
c16	0.365	0.532	0.196	0.377
c17	0.205	0.925	0.107	
c18	0.360	0.411	0.227	0.502
c19	0.221	0.894		

- MR1.知識分享
- MR3.知識獲取
- MR2.知識創新
- MR4. 刪除 (c12)

	MR1	MR3	MR2	MR4
SS loadings	4.264	3.574	3.070	0.882
Proportion Var	0.224	0.188	0.162	0.046
Cumulative Var	0.224	0.413	0.574	0.621

```
> efa$scores
```

	MR1	MR3	MR2	MR4
[1,]	1.402167479	0.187838477	-1.2537167527	1.279342543
[2,]	-0.419055681	0.394402505	0.3166798651	1.906799005
[3,]	1.085330960	0.107331779	0.2080530320	1.175347268
[4,]	-1.796821381	0.891926599	1.6657578761	0.753229570
[5,]	1.126557522	0.306904305	0.2187840659	0.837639378
[6,]	1.135976889	0.635736865	0.0949324673	0.173110056
...				
[200,]	0.536445634	-0.322323829	-1.2022886994	0.207571768

- 第一次因素分析:
 - 第四個因素只包含一個題項c12 (少於3 題), 因素無法命名。
 - 刪除此題項, 再進行一次因素分析。
- 原則: 逐次刪除最不適切的題項, 勿同時刪除多個變數。
 - 因素抽取方法: 主軸法 + 最大變異直交轉軸法

```
> Knowledge_Management_2 <- Knowledge_Management
> Knowledge_Management_2$c12 <- NULL
>
> check_factorstructure(Knowledge_Management_2)
# Is the data suitable for Factor Analysis?

- Sphericity: Bartlett's test of sphericity suggests that there is sufficient significant
correlation in the data for factor analysis (Chisq(153) = 3062.23, p < .001).
- KMO: The Kaiser, Meyer, Olkin (KMO) overall measure of sampling adequacy suggests that data seems
appropriate for factor analysis (KMO = 0.86). The individual KMO scores are: c1 (0.87), c2 (0.81), c3
(0.78), c4 (0.81), c5 (0.87), c6 (0.89), c7 (0.89), c8 (0.88), c9 (0.77), c10 (0.96), c11 (0.84), c13
(0.78), c14 (0.93), c15 (0.83), c16 (0.94), c17 (0.84), c18 (0.91), c19 (0.90).

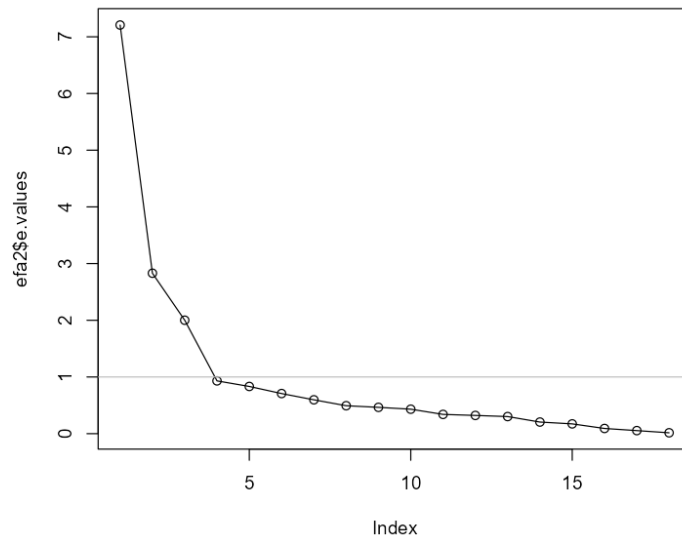
>
> efa2 <- fa(Knowledge_Management_2, nfactors = 3, rotate = "varimax",
+           fm = "pa" , scores = "regression")
```



```

> efa2$communality
      c1      c2      c3      c4      c5      c6
0.3743457 0.6192817 0.8439875 0.4291923 0.3976103 0.2188368
      c7      c8      c9      c10     c11     c13
0.5313825 0.6685175 0.8040428 0.6190435 0.5163039 0.8221319
      c14     c15     c16     c17     c18     c19
0.5889172 0.9793284 0.5454986 0.8698462 0.4563757 0.8046645
> efa2$e.values
 [1] 7.20781546 2.83003865 2.00145538 0.93109501 0.83278861
 [6] 0.70766152 0.59728803 0.49314380 0.46441950 0.43178212
[11] 0.33991157 0.32169304 0.30174530 0.20490705 0.17274668
[16] 0.09181834 0.05403028 0.01565966
    
```

陡坡圖



```
> 解說變異量 <- data.frame(特徵值 = efa2$e.values,
+                          變異數的百分比 = 100 * efa2$e.values/sum(efa2$e.values),
+                          累積的百分比 = 100 * cumsum(efa2$e.values/sum(efa2$e.values)))
>
> rownames(解說變異量) <- colnames(Knowledge_Management_2)
> round(解說變異量, 3)
```

	特徵值	變異數的百分比	累積的百分比
c1	7.208	40.043	40.043
c2	2.830	15.722	55.766
c3	2.001	11.119	66.885
c4	0.931	5.173	72.058
c5	0.833	4.627	76.684
c6	0.708	3.931	80.616
c7	0.597	3.318	83.934
c8	0.493	2.740	86.674
c9	0.464	2.580	89.254
c10	0.432	2.399	91.653
c11	0.340	1.888	93.541
c13	0.322	1.787	95.328
c14	0.302	1.676	97.005
c15	0.205	1.138	98.143
c16	0.173	0.960	99.103
c17	0.092	0.510	99.613
c18	0.054	0.300	99.913
c19	0.016	0.087	100.000

```
> efa2$loadings
```

Loadings:

	PA1	PA3	PA2
c1	0.102		0.602
c2			0.780
c3			0.913
c4			0.651
c5	0.126		0.610
c6	0.138	0.158	0.418
c7	0.639	0.325	0.134
c8	0.770	0.191	0.196
c9	0.874	0.180	
c10	0.754	0.226	
c11	0.662	0.167	0.225
c13	0.886	0.170	
c14	0.466	0.598	0.120
c15	0.190	0.963	0.122
c16	0.391	0.592	0.204
c17	0.183	0.910	
c18	0.404	0.485	0.241
c19	0.198	0.874	

	PA1	PA3	PA2
SS loadings	4.256	3.796	3.038
Proportion Var	0.236	0.211	0.169
Cumulative Var	0.236	0.447	0.616

輸出: 成份矩陣, 因素命名

```
> print.psych(fa(Knowledge_Management_2, nfactors = 3, rotate = "varimax",
+               fm = "pa" , scores = "regression"), cut = 0.35, sort = TRUE)
```

Factor Analysis using method = pa

Call: fa(r = Knowledge_Management_2, nfactors = 3, rotate = "varimax",
scores = "regression", fm = "pa")

Standardized loadings (pattern matrix) based upon correlation matrix

item	PA1	PA3	PA2	h2	u2	com
c13	12	0.89		0.82	0.178	1.1
c9	9	0.87		0.80	0.196	1.1
c8	8	0.77		0.67	0.331	1.3
c10	10	0.75		0.62	0.381	1.2
c11	11	0.66		0.52	0.484	1.4
c7	7	0.64		0.53	0.469	1.6
c15	14		0.96	0.98	0.021	1.1
c17	16		0.91	0.87	0.130	1.1
c19	18		0.87	0.80	0.195	1.1
c14	13	0.47	0.60	0.59	0.411	2.0
c16	15	0.39	0.59	0.55	0.455	2.0
c18	17	0.40	0.48	0.46	0.544	2.4
c3	3		0.91	0.84	0.156	1.0
c2	2		0.78	0.62	0.381	1.0
c4	4		0.65	0.43	0.571	1.0
c5	5		0.61	0.40	0.602	1.1
c1	1		0.60	0.37	0.626	1.1
c6	6		0.42	0.22	0.781	1.5

```
> efa2$scores
```

	PA1	PA3	PA2
[1,]	1.6334459478	0.44274542	-1.214100508
[2,]	-0.2575845038	0.81716005	0.416899657
...			
[199,]	0.3880896876	0.74198934	-0.187360676
[200,]	0.5370199233	-0.32758234	-1.089028733

```
> KM_FA_Scores <- efa2$scores
```

```
> colnames(KM_FA_Scores) <- c("知識分享", "知識獲取", "知識創新")
```

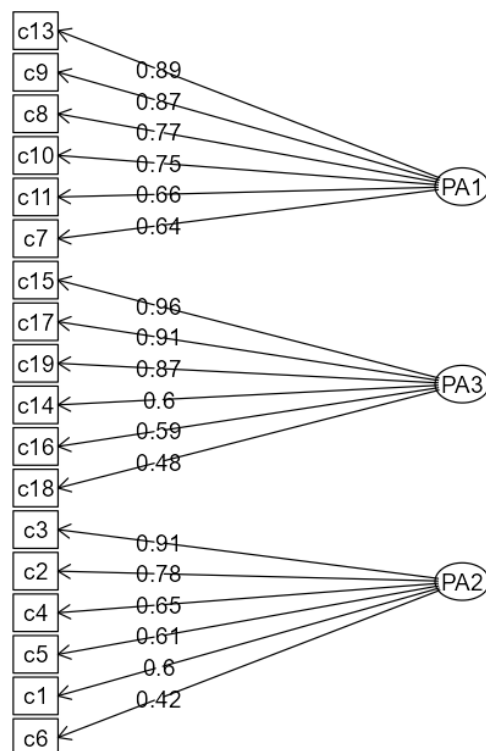
```
> head(KM_FA_Scores)
```

	知識分享	知識獲取	知識創新
[1,]	1.6334459	0.4427454	-1.2141005
[2,]	-0.2575845	0.8171600	0.4168997
[3,]	1.2922374	0.3579175	0.2415262
[4,]	-1.7103957	1.0025995	1.6880902
[5,]	1.3186540	0.4068162	0.2505948
[6,]	1.2794222	0.5000737	0.0657161

因素分析結果摘要表(I)

```
> fa.diagram(efa2, cut = 0.35, sort = TRUE, digits = 2, rsize = 0.5)
```

Factor Analysis



題項變數及題目	最大變異法直交轉軸後之因素負荷量			共同性
	知識分享	知識獲取	知識創新	
c13.本校教師會於同仁會議中分享其處理學生問題的策略。	.884	.171	.088	.818
c9.本校教師會在相關會議中提供意見供其他教師分享。	.871	.180	.092	.800
c8.本校教師會將班級經營的有效策略，與其他教師分享。	.772	.191	.195	.670
c10.本校行政事務處理流程有完整紀錄，以供同仁分享參考。	.755	.225	.009	.621
c11.本校教師很少於教學研討會上，分享其教學經驗。	.663	.166	.224	.517
c7.校長會積極鼓勵同仁，分享研習吸取的新知能。	.639	.325	.133	.532
c15.學校會鼓勵教師透過教學觀摩，以獲取專業知能。	.189	.967	.122	.986
c17.學校鼓勵教師透過教師社群活動，以獲取專業知能。	.184	.909	.089	.868
c19.學校會影印相關教育新知給教師，以增進教師知能。	.199	.873	.037	.802
c14.學校鼓勵同仁參訪標竿學校以獲取教學及行政知能。	.466	.598	.119	.589
c16.學校積極鼓勵教師參與研習活動，以獲取專業知能。	.392	.592	.204	.545
c18.學校鼓勵教師透過數位資料來獲取新知識。	.405	.485	.240	.456
c3.教師會積極地在其負責的行政工作上創新展現。	.051	.084	.919	.854
c2.本校教師會積極尋求班級經營上的創新。	.070	.080	.779	.618
c4.本校教師會應用研習心得於教育品質的提升。	.070	.036	.649	.427
c5.本校會激勵教師以創新理念提升學生學習成效。	.127	.095	.609	.396
c1.本校常鼓勵教師創新教學或工作創新。	.102	.035	.602	.374
c6.本校鼓勵教師以創新有效方法激勵學生學習動機。	.138	.158	.418	.219
特徵值	4.257	3.796	3.039	11.092
解釋變異量%	23.648	21.091	16.884	61.623
累積解釋變異量%	23.648	44.739	61.623	

因素分析結果摘要表(II)

題項變數及題目	直接斜交法斜交轉軸後之結構矩陣			共同性
	知識分享	知識獲取	知識創新	
c13.本校教師會於同仁會議中分享其處理學生問題的策略。	.907	.209	.406	.824
c9.本校教師會在相關會議中提供意見供其他教師分享。	.901	.212	.412	.813
c8.本校教師會將班級經營的有效策略，與其他教師分享。	.852	.307	.401	.732
c10.本校行政事務處理流程有完整記錄，以供同仁分享參考。	.825	.114	.414	.696
c11.本校教師很少於教學研討會上，分享其教學經驗。	.767	.330	.349	.605
c7.校長會積極鼓勵同仁，分享研習吸取的新知能。	.750	.240	.515	.594
c3.教師會積極的在其負責的行政工作上創新展現。	.192	.896	.197	.806
c2.本校教師會積極尋求班級經營上的創新。	.188	.822	.182	.676
c4.本校教師會應用研習心得於教育品質的提升。	.158	.738	.119	.548
c5.本校會激勵教師以創新理念提升學生學習成效。	.226	.711	.195	.507
c1.本校常鼓勵教師創新教學或工作創新。	.194	.694	.114	.485
c6.本校鼓勵教師以創新有效方法激勵學生學習動機。	.228	.530	.252	.299
c15.學校會鼓勵教師透過教學觀摩，以獲取專業知能。	.407	.237	.966	.939
c17.學校鼓勵教師透過教師社群活動，以獲取專業知能。	.382	.199	.935	.880
c19.學校會影印相關教育新知給教師，以增進教師知能。	.378	.146	.914	.844
c14.學校鼓勵同仁參訪標竿學校以獲取教學及行政知能。	.607	.229	.760	.653
c16.學校積極鼓勵教師參與研習活動，以獲取專業知能。	.541	.312	.752	.618
c18.學校鼓勵教師透過數位資料來獲取新知識。	.545	.348	.656	.521
累積解釋變異量%				66.885

因素分析實例

表 4-14 收藏行為之因素分析

因素	題號	因素負荷量	特徵值	解釋變異量
收藏意願	A3	0.78	2.05	20.54%
	A2	0.77		
	A6	0.63		
	A1	0.58		
收藏完整性	A4	0.72	1.99	19.89%
	A7	0.70		
	A8	0.67		
	A5	0.66		
收藏增值	A10	0.92	1.96	19.62%
	A11	0.91		

KMO = 0.62
Bartlett 球形檢定=636.10，自由度= 45，Sig=0.00
總解釋變異量為 60.04%

表 4-15 從眾行為之因素分析

因素	題號	因素負荷量	特徵值	解釋變異量
從眾傾向	B2	0.79	1.70	56.53%
	B1	0.75		
	B3	0.72		

KMO = 0.64
Bartlett 球形檢定=146.79，自由度= 3，Sig=0.00
總解釋變異量為 56.53%

表 4-16 參考群體之因素分析

因素	題號	因素負荷量	特徵值	解釋變異量
參考群體依 賴度	C2	0.88	2.23	74.20%
	C3	0.87		
	C1	0.83		

KMO = 0.71
Bartlett 球形檢定=497.73，自由度= 3，Sig=0.00
總解釋變異量為 74.20%

表 4-17 消費體驗之因素分析

因素	題號	因素負荷量	特徵值	解釋變異量
感官體驗	D1	0.84	2.44	34.91%
	D2	0.81		
	D5	0.61		
	D4	0.60		
	D3	0.52		
情感體驗	D6	0.84	1.80	25.73%
	D7	0.82		

KMO = 0.62
Bartlett 球形檢定=636.10，自由度= 45，Sig=0.00
總解釋變異量為 60.64%

表 4-18 顧客價值之因素分析

因素	題號	因素負荷量	特徵值	解釋變異量
新奇價值	E12	0.88	2.21	18.42%
	E13	0.76		
	E11	0.75		
視覺價值	E4	0.79	2.05	17.10%
	E5	0.75		
	E6	0.73		
購買成本	E1	0.81	1.75	14.56%
	E2	0.72		
	E3	0.63		
社會價值	E7	0.81	1.74	14.50%
	E8	0.76		
	E9	0.50		

KMO = 0.76
Bartlett 球形檢定=1463.17，自由度= 66，Sig=0.00
總解釋變異量為 64.59%

- 因素分析完，要進行**量表各層面**與**總量表**的信度考驗。
- 信度: 指量表工具所測得結果的**穩定性(stability)**及**一致性(consistency)**。
- 信度愈大，測量標準誤愈小。

某一測量工具，測得一變異量值為： V_0 則 $V_0 = V_t + V_e$

真實變異量與觀察變異量之比，即為信度。

$$\text{信度} = \frac{V_t}{V_0} = \frac{V_0 - V_e}{V_0} = 1 - \frac{V_e}{V_0}$$

- V_0 : 觀察值
- V_t : 真實值
- V_e : 誤差值

因 V_t 難直接估計，故經公式轉換，

$$\text{信度} = 1 - \frac{\text{誤差變異量}}{\text{觀察變異量}}$$

- 內部一致性信度：Cronbach alpha，適用李克特式(Likert scale)量表。

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum S_i^2}{S^2} \right)$$

$$\text{信度} = 1 - \frac{\text{誤差變異量}}{\text{觀察變異量}}$$

1. K : 量表總題數。
2. $\sum_i S_i^2$: 量表題項總變異和。
3. S : 量表題項加總後變異數。

- 題項數愈多， $K/(K-1)$ 愈接近於1， $\sum_i S_i^2 / S^2$ 的值愈接近0，alpha 愈接近1。
- 題目間相關係數愈大，alpha值愈大。

Cronbach alpha之計算

樣本	量表試題								總分
	01	02	03	04	05	06	07	08	
A	5	1	2	5	2	5	4	3	27
B	5	1	2	4	3	5	5	2	27
C	5	2	2	5	3	5	5	2	29
D	5	1	2	5	3	5	5	3	29
E	5	1	2	5	3	5	5	2	28
F	4	1	1	5	3	4	4	2	24

題號	敘述統計					
	個數	最小值	最大值	平均數	標準差	變異數
A1	6	4	5	4.833	0.408	0.167
A2	6	1	2	1.167	0.408	0.167
A3	6	1	2	1.833	0.408	0.167
A4	6	4	5	4.833	0.408	0.167
A5	6	2	3	2.833	0.408	0.167
A6	6	4	5	4.833	0.408	0.167
A7	6	4	5	4.667	0.516	0.267
A8	6	2	3	2.333	0.516	0.267
ΣS ²						1.533
總分	6	24	29	27.333	1.862	3.467

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum S_i^2}{S^2} \right) = \frac{8}{8-1} \left(1 - \frac{1.533}{3.467} \right) = \frac{8}{7} (0.558) = 0.637$$

樣本	量表試題								Z總分
	ZA1	ZA2	ZA3	ZA4	ZA5	ZA6	ZA7	ZA8	
A	0.408	-0.408	0.408	0.408	-2.041	0.408	-1.291	1.291	-0.816
B	0.408	-0.408	0.408	-2.041	0.408	0.408	0.645	-0.645	-0.816
C	0.408	2.041	0.408	0.408	0.408	0.408	0.645	-0.645	0.408
D	0.408	-0.408	0.408	0.408	0.408	0.408	0.645	1.291	3.569
E	0.408	-0.408	0.408	0.408	0.408	0.408	0.645	-0.645	1.633
F	-2.041	-0.408	-2.041	0.408	0.408	-2.041	-1.291	-0.645	-7.652

	個數	最小值	最大值	平均數	標準差	變異數
ZA1	6	-2.041	.408	.00000	1.000000	1.000
ZA2	6	-.408	2.041	.00000	1.000000	1.000
ZA3	6	-2.041	.408	.00000	1.000000	1.000
ZA4	6	-2.041	.408	.00000	1.000000	1.000
ZA5	6	-2.041	.408	.00000	1.000000	1.000
ZA6	6	-2.041	.408	.00000	1.000000	1.000
ZA7	6	-1.291	.645	.00000	1.000000	1.000
ZA8	6	-.645	1.291	.00000	1.000000	1.000
Z總分	6	-7.652	4.082	.00000	4.288601	18.392

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum S_i^2}{S^2} \right) = \frac{8}{8-1} \left(1 - \frac{8}{18.392} \right) = \frac{8}{7} (0.565) = 0.646$$

- 信度是指: 測驗得到結果的一致性或穩定性，而非測驗本身。要說「此測驗分數是可信賴的」，不是「測驗是可信賴的」。
- 信度可能因不同時間，不同受試者而出現不同結果。
- 信度考驗完全是依據統計方法而求得。
- 信度: 為真實分數的變異數與觀察分數的變異數之比例。
- 使用前人編製之量表，需再經預試，以重新考驗其信度。
 - 最好同時提供信度係數之信賴區間。

信度的種類

- **外在信度 (external reliability)**
 - 不同時間測量時，量表一致性的程度。
 - **再測信度 (test-retest reliability)**：以同一測量工具實施二次測量結果的相關程度（即相關係數）來估量，相關程度愈高，表示測量結果愈穩定，亦即信度愈高；反之則反。
- **內在信度 (internal reliability)**
 - 量表是否測量單一概念
 - **折半信度 (split-half reliability)**：將測驗結果之題目平均分為兩組（通常以題號為準，單號題一組，雙號題一組），分別計算受試者在各組的得分，並進一步求得這兩組分數的相關程度，然後依據「斯 - 布公式」(Spearman-Brown Formula)計算，所得結果即為信度係數。
 - 內部一致性信度：Cronbach alpha
 - 庫李信度
 - RK信度
 - ...

- 信度是效度的必要條件，非充份條件。
- 信度低，效度一定低; 信度高，未必表示效度高。
- 有效的測驗必須是可信的測驗。
- 可信的的測驗未必是有效的測驗。

內部一致性信度係數值	層面或構念	整個量表
α 係數 $<.50$	不理想，捨棄不用	非常不理想，捨棄不用
$.50 \leq \alpha$ 係數 $<.60$	可以接受，增列題項或修改語句	不理想，重新編製或修訂
$.60 \leq \alpha$ 係數 $<.70$	尚佳	勉強接受，最好增列題項或修改語句
$.70 \leq \alpha$ 係數 $<.80$	佳（信度高）	可以接受
$.80 \leq \alpha$ 係數 $<.90$	理想（甚佳，信度很高）	佳（信度高）
α 係數 $\geq .90$	非常理想（信度非常好）	非常理想（甚佳，信度很高）



信度分析實例操作： 求「知識創新」層面的信度

- 量表: 「學校知識管理量表」
- 已進行「項目分析」與「因素分析」:
 - 三個因素構念:
 - 知識創新: c1, c2, c3, c4, c5, c6
 - 知識分享: c7, c8, c9, c10, c11, c13
 - 知識獲取: c14, c15, c16, c17, c18, c19
- 求此三個因素(構念)及量表測驗結果的信度。

學校知識管理量表

題項	完全不符合	多數不符合	半數不符合	多數符合	完全符合
01.本校常鼓勵教師創新教學或工作創新。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02.本校教師會積極尋求班級經營上的創新。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
03.教師會積極的在其負責的行政工作上創新展現。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
04.本校教師會應用研習心得於教育品質的提升。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
05.本校會激勵教師以創新理念提升學生學習成效。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
06.本校鼓勵教師以創新有效方法激勵學生學習動機。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
07.校長會積極鼓勵同仁, 分享研習吸取的新知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
08.本校教師會將班級經營的有效策略, 與其他教師分享。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
09.本校教師會在相關會議中提供意見供其他教師分享。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.本校行政事務處理流程有完整紀錄, 以供同仁分享參考。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.本校教師很少於教學研討會上, 分享其教學經驗。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12.本校同仁會於朝會上分享其研習的心得與知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13.本校教師會於同仁會議中分享其處理學生問題的策略。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14.學校鼓勵同仁參訪標竿學校以獲取教學及行政知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15.學校會鼓勵教師透過教學觀摩, 以獲取專業知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.學校積極鼓勵教師參與研習活動, 以獲取專業知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17.學校鼓勵教師透過教師社群活動, 以獲取專業知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
18.學校鼓勵教師透過數位化資料來獲取新知識。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19.學校會影印相關教育新知給教師, 以增師知能。.....	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

```
> Knowledge_Management <- read.csv("data/學校知識管理_1.csv")
> head(Knowledge_Management)
  c1 c2 c3 c4 c5 c6 c7 c8 c9 c10 c11 c12 c13 c14 c15 c16 c17 c18 c19
1  4  3  3  3  3  4  4  4  4  5  5  2  5  4  4  4  3  5  3
2  4  4  4  4  4  5  4  3  3  3  4  2  3  4  4  4  3  5  3
3  4  4  4  4  4  4  4  4  4  4  5  2  5  4  4  4  3  5  3
4  5  5  5  5  4  4  2  2  2  2  3  2  2  4  4  4  4  5  3
5  4  4  4  4  4  5  4  4  4  4  5  2  5  4  4  4  4  5  3
6  4  4  4  4  4  4  4  4  4  4  5  3  5  4  4  4  4  5  5
> dim(Knowledge_Management)
[1] 200 19
> knowledge_creative <- Knowledge_Management[, 1:6]
```



總量表及分量表的信度分析操作

```
> library(psych)
> reliability_output <- psych::alpha(knowledge_creative, check.keys = T)
> reliability_output
```

Reliability analysis

Call: psych::alpha(x = knowledge_creative, check.keys = T)

```
raw_alpha std.alpha G6(smc) average_r S/N ase mean sd median_r
0.83      0.83      0.83      0.45 4.9 0.018 3.9 0.51 0.46
```

95% confidence boundaries

```
lower alpha upper
Feldt      0.79 0.83 0.87
Duhachek 0.80 0.83 0.87
```

Reliability if an item is dropped:

```
raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
c1      0.81      0.81      0.81      0.46 4.3 0.021 0.025 0.46
c2      0.78      0.78      0.77      0.42 3.6 0.024 0.019 0.37
c3      0.76      0.76      0.74      0.38 3.1 0.027 0.013 0.32
c4      0.81      0.80      0.80      0.45 4.1 0.022 0.027 0.42
c5      0.81      0.81      0.80      0.46 4.2 0.021 0.028 0.43
c6      0.84      0.84      0.84      0.52 5.3 0.018 0.019 0.52
```

Item statistics

```
n raw.r std.r r.cor r.drop mean sd
c1 200 0.69 0.70 0.61 0.55 4.1 0.65
c2 200 0.81 0.80 0.79 0.70 3.9 0.72
c3 200 0.88 0.88 0.90 0.81 3.9 0.72
c4 200 0.73 0.73 0.66 0.60 3.8 0.65
c5 200 0.72 0.71 0.63 0.57 3.5 0.72
c6 200 0.57 0.58 0.43 0.40 4.2 0.66
```

Non missing response frequency for each item

```
1 2 3 4 5 miss
c1 0 0.01 0.15 0.60 0.24 0
c2 0 0.01 0.31 0.50 0.18 0
c3 0 0.00 0.26 0.54 0.20 0
c4 0 0.01 0.32 0.56 0.10 0
c5 0 0.06 0.42 0.44 0.08 0
c6 0 0.01 0.10 0.58 0.30 0
```

- 求「**知識分享**」層面(分量表)的信度
 - Cronbach's alpha=0.912; Std Cronbach's alpha=0.914
 - 觀察「項目刪除後之Cronbach's alpha值」。
 - 「**知識分享**」層面的內部一致性信度非常理想。

- 求「**知識獲取**」層面(分量表)的信度
 - Cronbach's alpha=0.915; Std Cronbach's alpha=0.916
 - 觀察「項目刪除後之Cronbach's alpha值」。(若刪除c18: alpha=0.921)
 - 「**知識獲取**」層面的內部一致性信度非常理想。

- 求量表「**學校知識管理量表**」(總量表)的信度
 - Cronbach's alpha=0.910; Std Cronbach's alpha=0.905
 - 量表的信度佳。

- 「再測信度」操作(略)。