

變項中心化與多元共線性的玄機

李茂能

嘉義大學

中文摘要

變項中心化常可令截距的解釋具有實質意義與減輕多元共線性的威脅，但對於交互作用項中心化的議題則尚有爭論。過去研究發現，有時交互作用項的中心化並無法有效改善多元共線性現象。為探究此矛盾現象，本文乃透過理論的推導與實徵資料的分析，發現迴歸係數變異量涉及三大要素：估計誤差變異量、預測變項之變異量、變項間之相關性。交互作用項會出現多元共線性現象，乃是這三大主因交織而成，並非單一因素所致。過去對於多元共線性的研究，常侷限於變項間之相關性上，因而導致研究結論不一致的現象。

多元共線性的診斷不能單靠 VIF/Tolerance 指標，其它指標如 CI/Variance proportion/Eigenvalue 指標亦須同時納入診斷之，才能正確診斷出來。文中除公式推導之外，並以模擬實例闡釋了 VIF/Tolerance 指標旨在反映預測變項間之相關程度，而 CI/Variance proportion/Eigenvalue 指標則在反映預測變項的離散量，這兩類指標反映多元共線性的不同面向的資料結構。

文中並提出預測變項「變異比差異」理論說明，為何變項中心化有時可以改善多元共線性的威脅，但有時卻不能改善的矛盾現象。

關鍵詞：變項中心化、多元共線性、交互作用項、變異比

Variable Mean-Centering and Multicollinearity

Mao-Neng Li

National Chia-Yi University

Abstract

To reduce multicollinearity and to make the interpretation of the intercept meaningful, it is frequently recommended that an interaction variable be based on mean-centered variables. Previous studies however have found that this approach may be futile. The variance of a regression coefficient is affected by three factors: error variance, variance of the predicting variables, and the correlation between predictors.

Multicollinearity for the interaction term can only be detected correctly by analyzing simultaneously each of these three inter-related factors. Previous researchers have only focused on the correlation between predictors by relying on the VIF/Tolerance index to detect multicollinearity. Consequently, inconsistent findings regarding the merits of mean-centering have been reported. Using an illustrated example, we show that the Condition/Variance proportion Index which reflects the variance of the predictors in addition to VIF be examined to effectively detect multicollinearity.

Interestingly, mean-centering the predicting variables for reducing the threat of multicollinearity can be valid only under the condition that there is not a significant increased difference of variance ratio (SD/Mean) especially between the variance ratio of the pre-centering interaction variable and the variance ratio of the post-centering interaction term.

Keywords: Mean-Centering, Multicollinearity, Interaction Term, Variance Ratio

壹、序言

交互作用項中心化 (Centering, 亦可稱為變項置中) 在多元迴歸分析、結構方程模式、成長模式與階層線性模式的論文中, 常被用來解決多元共線性的問題或使得截距的解釋具有實質的意義。交互作用項在 HLM (Hierarchical Linear Model) 中可能發生在一階、二階及跨層次的交互作用上, 因此 HLM 軟體就內建有變項中心化功能。足見交互作用項的中心化是量化研究者必須深入理解的重要課題。

不過, 到底變項中心化可否有效消除或減輕多元共線性的現象, 卻仍有爭論之處。例如: Belsley (1984), Gatignon & Vosgerau (2005) 認為交互作用項的平均數置中, 雖可降低單一變項與交互作用項間之線性相關, 但仍無法解決或減輕多元共線性的威脅。交互作用項的多元共線性問題並非資料結構不良 (變項間缺乏獨立性) 所致, 而是加入交互作用項所產生的模式界定問題。中心化使得原來雙變項的共線性變成多變項的共線性, 其來源改變了, 但共線性大小並未改變。Echambadi & Hess (2007) 也證明平均數置中徒勞無功, 並無法解決多元共線性問題。該文證實了中心化變項既無法改善參數估計值之運算精確性, 亦無法提高整體 R^2 之解釋力 (即使 VIF (variance inflation factor) 下降); 因為交乘積矩陣 ($X'X$) 的行列式值在中心化與未中心化的資料上均完全相同。可見交互作用項的共線性問題並未有任何的改善, 這些論點似乎打破了過去傳統之看法, 似乎仍有其它重要因素也會導致多元共線性問題。另外, Hofmann (2007) 認為錯誤的中心化決定會導致你所考驗的理論模式與提出的假設不一致; 而中心化方法選擇的錯誤亦會導致虛假的跨層次交互作用 (錯誤顯示 level-2 的變項可以預測 level-1 的斜率)。

因此, 中心化的議題、多元共線性的迷思以及兩者之矛盾關係乃是統計分析者的重要課題, 值得進一步深入分析與釐清。以下, 首先釐清中心化的意義、用途與類別; 接著, 簡述多元共線性的意義與症狀, 並推導其原因。最後, 提出解決之道。

貳、中心化的定義與用途

中心化係指原點 (zero point) 的平移, 常將原始分數減去總平均數轉換成離差分數。教育或社會科學研究中, 心理構念常缺乏一個可解釋而有意義的原點; 例如: 如果使用原始量尺: $HDL (好膽固醇) = -24.990 + 0.498LDL (壞膽固醇) + 2.459BMI - 0.019 (LDL * BMI)$, 當預測值為 0 時, 所獲得之截距為負值, 不具有實質之應用性。如果將這些預測變項加以置中後, 其預測公式變為: $HDL (好膽固醇) = 47.555 + 0.080 (LDL - 215) - 1.537 (BMI - 22.5) - 0.019 (LDL - 215) (BMI - 22.5)$ 。因此當預測值為 0 時, 所獲得之截距為正值, 就具有實質之意義; 又如, LDL & BMI 為平均值時, 預估之好膽固醇平均值為 47.555, 都具有應用上之實質意義。因此, 需要將預測變項平移至有意義的新原點, 以便於解釋, 注意中心化後各觀察值的相對位置並不會改變。又如在縱貫式的研究上, 中心化目的在找到一個時間點, 對於截距的解釋才具有意義或才能有效回答待答問題。不同的中心化方法, 其所得的係數與截距將有不同的意義。當然, 假如預測變項值為具有實質意義的 0, 通常就不必進行變項中心化了。

參、中心化的類別與運用

變項中心化的類別可以分為三類:

一、殘差中心化 (Residual-centering)

Lance (1988), Little, Bovaird & Widaman (2006) 之正交化乘積指標法，包含兩大步驟：

(一) 首先，須先建立未中心化指標之交乘積項，利用這些交乘積項當作效標進行迴歸分析；將殘差存檔以作後續分析之資料，

(二) 建立潛在交互作用項 ($\xi_1\xi_2$)，並以交乘積殘差項 (Res11~Res22) 作為其指標 (參見圖 1)。

新建立的正交化交互作用項 (orthogonalized product terms)，僅包含代表交互作用效果的獨特變異量，因此交乘積或高次方項與主要效果間之關係完全獨立無關，而平均數中心化 (Mean-centering) 僅能降低兩者間之強度關係，此係殘差中心化之最大特色。

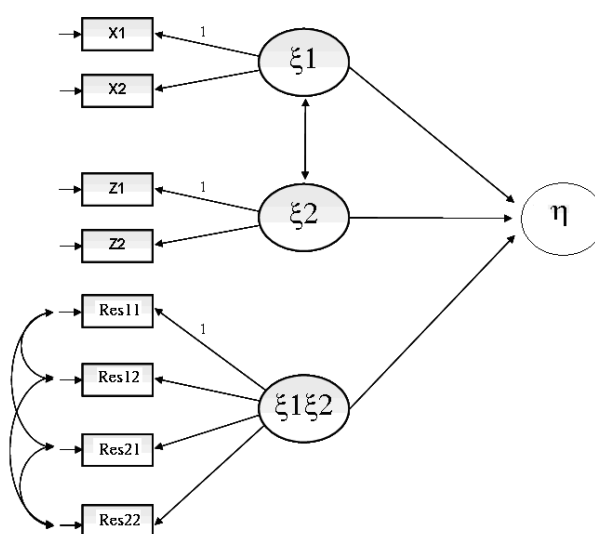


圖1 殘差中心化範例

二、平均數中心化 (Mean-centering)

平均數中心化常見於迴歸分析、SEM與HLM分析中，尤其是HLM多層次分析時，變項中心化更是無法逃避的重要課題。茲將常用的平均數中心化方法簡介如下，以便利研究者查考：

(一) 組平均數中心化

組平均數中心化 (Centering within clusters；又稱為 Group-Mean centering)，或稱為脈絡中心化 (Context centering，簡稱為 CWC)，適用時機有：

1、研究者旨在探究「層次一」中變項 X & Y 間之關係 (個體效果)，希望：(1) 獲得淨組內迴歸係數的估計值 (可排除組間變異量)，與 (2) 較精確估計斜率變異數。不過，當 $\beta_b = \beta_w$ 則要使用 CGM (Centering at the grand mean)，最為有效率。

2、研究者希望進行跨層次之交互作用 (Slopes as outcome)，以獲得純淨的「層次一」斜率 (此斜率受到其它預測變項的調節)，因為本法可以分離 cross-level & between group interaction 的效果 (Hofmann & Gavin, 1998)。

3、研究者旨在探究「層次一」變項間之交互作用。

(二) 總平均數中心化

總平均數中心化 (Centering at the grand mean, 又稱為 Grand-Mean centering, 簡稱為 CGM), 又稱為總平減置中, 適用時機為:

1、研究者旨在探究「層次二」中預測變項之淨效果 (以排除層次一共變項之效果)。

2、研究者旨在探究「層次二」中變項間的交互作用。

如欲探究變項在「層次一與層次二」是否具有差異效果 (亦即脈絡效果), Wu & Wooldridge (2005) 建議使用原始分數或使用 CGM (請在層次二中加入組平均數預測變項), 不必進行線性轉換即可區分個體層次效果與組織層次效果。

另外, 總平減置中與原始分數間之關係非常密切 (二者係等同模式), 所估計之參數可以透過線性轉換互轉, 細節請看下列數學公式一~公式三之推導。

$$\begin{aligned}
 y &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 \\
 y &= \gamma_0 + \gamma_1 X_1^c + \gamma_2 X_2^c + \gamma_3 X_1^c X_2^c \quad (X_1^c = X_1 - \bar{X}_1, X_2^c = X_2 - \bar{X}_2) \\
 y &= \gamma_0 + \gamma_1 (X_1 - \bar{X}_1) + \gamma_2 (X_2 - \bar{X}_2) + \gamma_3 (X_1 - \bar{X}_1)(X_2 - \bar{X}_2) \quad (\text{公式一}) \\
 &= \gamma_0 + (\gamma_1 - \gamma_3 \bar{X}_2) X_1 + (\gamma_2 - \gamma_3 \bar{X}_1) X_2 + \gamma_3 X_1 X_2 \\
 &= (\gamma_0 + \gamma_3 \bar{X}_1 \bar{X}_2 - (\gamma_1 \bar{X}_1 + \gamma_2 \bar{X}_2)) + (\gamma_1 - \gamma_3 \bar{X}_2) X_1 + (\gamma_2 - \gamma_3 \bar{X}_1) X_2 + \gamma_3 X_1 X_2
 \end{aligned}$$

由上式知, 總平減置中對於最高階交互作用項係數完全無影響 ($\beta_3 = \gamma_3$)

置中後 X_1 & X_2 的斜率為:

$$\begin{aligned}
 \beta_1 &= (\gamma_1 - \gamma_3 \bar{X}_2) \\
 \rightarrow \gamma_1 &= \beta_1 + \gamma_3 \bar{X}_2 \\
 \beta_2 &= (\gamma_2 - \gamma_3 \bar{X}_1) \\
 \rightarrow \gamma_2 &= \beta_2 + \gamma_3 \bar{X}_1
 \end{aligned} \quad (\text{公式二})$$

至於置中後截距為:

$$\begin{aligned}
 &\text{將上式變項} \\
 \beta_0 &= (\gamma_0 + \gamma_3 \bar{X}_1 \bar{X}_2 - (\gamma_1 \bar{X}_1 + \gamma_2 \bar{X}_2)) \\
 &\text{移項後可得;} \quad (\text{公式三}) \\
 \gamma_0 &= (\beta_0 + (\gamma_1 \bar{X}_1 + \gamma_2 \bar{X}_2) - \gamma_3 \bar{X}_1 \bar{X}_2)
 \end{aligned}$$

三、固定值中心化

固定值中心化 (Fixed-value centering) 之適用時機為: 成長變化分析 (Latent change analysis)。

以下列成長變化模式為例: $Y_{ij} = \pi_{0i} + \pi_{1i} a_{ti} + \varepsilon_{ij}$, 假設觀察兒童的年齡分別在 2,3,4,5,6 歲, 截距 π_{0i} 即當 $a_{ti}=0$ 時的預測值。變項置中會影響截距的意義, 例如:

如果研究者定義 2 歲為 $a_{ti}-2=0$, 作為資料蒐集的初始狀態, 如果研究者想了解研究中點 4 歲作為初始狀態, 則使用 $a_{ti}-4$, 如果研究者想了解研究終點 6 歲作為

初始狀態,則使用 ati-6。

以上係對於不同感興趣時間點的歸零設定方法，中心化目的在找到一個時間點，對於截距的解釋才具有意義或才能有效回答待答問題。

肆、多元共線性的意義

多元共線性 (multicollinearity) 最常出現於多元迴歸分析模式 (或 HLM 分析) 上，尤其當兩個或兩個以上之預測變項間具有高度相關或建置交互作用項時，常會導致個別參數之估計值不穩定及其標準誤過大之現象：觀察值如有些許變動就會導致估計參數的巨大變動。茲以下述迴歸模式 Y 與 X 矩陣說明之：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_k X_{ki} + \varepsilon_i \quad (\text{公式四})$$

$$X = \begin{bmatrix} 1 & X_{11} & \cdots & X_{k1} \\ \vdots & \vdots & & \vdots \\ 1 & X_{1N} & \cdots & X_{kN} \end{bmatrix},$$

式中第一欄位代表截距，其餘為預測變項；

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}; \varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}.$$

公式四可以矩陣方式表示之： $Y = X\beta + \varepsilon$ ，

如果 X 為方陣具有反矩陣，則可直接求解， $\beta = X^{-1}Y$ ；但一般 X 均為非方陣，因此在迴歸分析中，一般均需利用 $\beta = (XX)^{-1}XY$ 求解。假如 X 矩陣為完全線性相依矩陣，此時 (XX) 的行列式值等於 0，因而反矩陣 $(XX)^{-1}$ 將不存在，亦即當有任何兩個預測變項間具完全相關，或是一個預測變項係其它預測變項間之線性組合時，將無法利用 OLS 解出 Beta 估計值；此時，統計分析軟體的報表通常會出現「Matrix is singular」的警語。茲以表 1 的資料（含有截距、X1、X2 & X3）為例，說明如下：

表1 完全線性相依之原始資料

1	3	4	8
1	4	3	6
1	3	3	6
1	2	2	4
1	3	6	12

本資料集中，X2 & X3 具有完全正相關，X3 為 X2 的兩倍大。利用 SPSS 矩陣

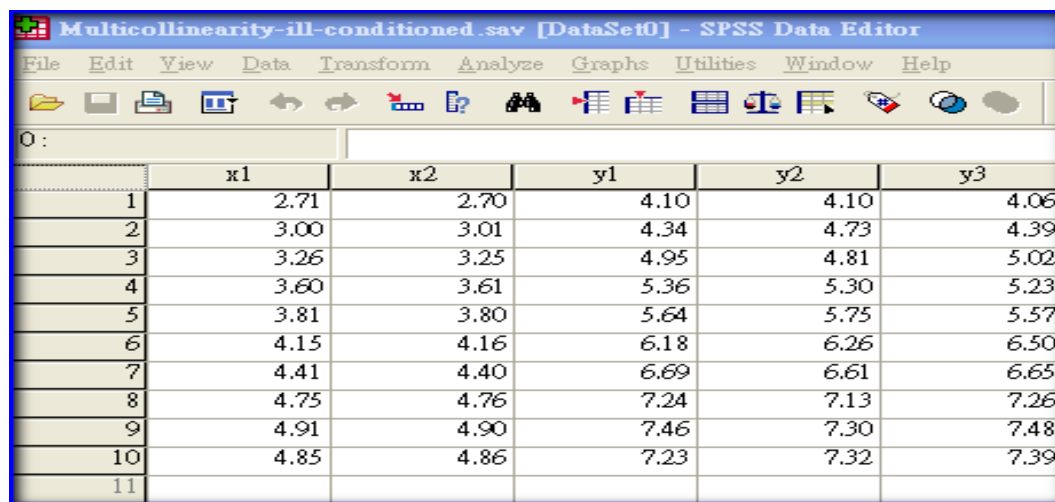
語言，不管資料是否中心化，求得之 X 行列式值均為 0，參見圖 2 之 SPSS 程式設計內容。

```
Matrix.
Compute a={1,3,4,8;1,4,3,6;1,3,3,6;1,2,2,4;1,3,6,12}.
Print a/title 'Raw data'.
Compute m1={t(a)*a}.
Print m1/title 'XX for uncentered data'.
compute f=ginv(m1).
Print f/title 'generalized inverse for singular XX--Raw score'.
Compute d=det(m1).
Print d/title 'Determinmant for Raw score'.
Compute b={1,0,0.4,.8;1,1,-0.6,-1.2;1,0,-0.6,-1.2;1,-1,-1.6,-3.2;1,0,2.4,4.8}.
print b/title 'Mean-centered data'.
compute m2=(t(b)*b).
print m2/ title 'XX for centered data'.
compute e=det(m2).
print e/title "Determinant for Mean-centered Data".
compute h=ginv(m2).
Print h/title 'generalized inverse for singular XX--mean-centered data'.
end matrix.
```

圖2 SPSS矩陣語言程式設計（中心化與未中心化）

就數學用語而言，就是 X 矩陣的 Rank 必須等於參數數目，如果小於參數數目，反矩陣值就不存在，亦即 $\text{Rank}(X'X) = \text{Rank}(X) = \text{參數數目}$ ， $(X'X)^{-1}$ 值存在。反之， $\text{Rank}(X'X)$ 小於參數數目，則 $(X'X)^{-1}$ 值不存在。

至於當 (XX) 的行列式值逼近於 0 時 (near-perfect multicollinearity)，會產生矩陣結構不良「ill-conditioned matrix」現象，此種現象在因素分析時亦常出現。茲以圖 3 資料為例，注意本資料的效標變項 y_1, y_2 & y_3 之數據僅有些微之不同，但其 Beta 估計值卻有顯著之不同；換言之，參數估計值會因資料些微之變動而產生重大改變，參數估計值呈現非常不穩定之現象，請參見表 2 三個迴歸模式中各參數之不穩定估計值與巨大 VIF 值。



	x1	x2	y1	y2	y3
1	2.71	2.70	4.10	4.10	4.06
2	3.00	3.01	4.34	4.73	4.39
3	3.26	3.25	4.95	4.81	5.02
4	3.60	3.61	5.36	5.30	5.23
5	3.81	3.80	5.64	5.75	5.57
6	4.15	4.16	6.18	6.26	6.50
7	4.41	4.40	6.69	6.61	6.65
8	4.75	4.76	7.24	7.13	7.26
9	4.91	4.90	7.46	7.30	7.48
10	4.85	4.86	7.23	7.32	7.39
11					

圖 3 矩陣結構不良之資料

圖 3 資料中， x_1 & x_2 之相關係數 $r_{x_1x_2}=.999915$ ，而以 y_1, y_2, y_3 當效標， x_1, x_2 為預測變項的決定係數分別為： $R^2_1=.997$ ， $R^2_2=.995$ ， $R^2_3=.990$ 。由表 2 之信賴區間大都包含 0 得知，除了 y_1 效標之外， x_1 & x_2 之迴歸係數在 y_2 & y_3 上均未達 .05 之顯著水準（反映個別預測變數沒有獨特預測力），但其相關之總體預測力 R^2 卻具有顯著之高預測力（.990~.997），因而出現矛盾之現象，此即典型「ill-conditioned matrix」之徵兆，VIF 值異常的高（5868.654）， x_1 & x_2 參數估計值（beta）會呈現非常不穩定之現象；例如： x_1 的估計值在三個迴歸模式上分別為 5.897，-2.002 & 1.526， x_2 的估計值在三個迴歸模式上分別為 -4.341，3.461 & .061。雖然效標變項 y_1 ， y_2 & y_3 之數據僅有些微變動，但相關之迴歸係數卻出現非常不穩定之變異差距。

表 2 矩陣結構不良之分析結果：不穩定的參數估計值

Model	Unstandardized Coefficients 95% Confidence interval for B Collinearity Statistic					
	B	Std. Error	Lower Bound	Upper bound	Tolerance	VIF
(constant)	-0.209	0.128	-0.511	0.930		
x1	5.897	2.409	0.200	11.593	0.000	5868.654
x2	-4.341	2.404	-10.025	1.342	0.000	5868.654
a. Dependent Variable : y1						
Model	Unstandardized Coefficients 95% Confidence interval for B Collinearity Statistic					
	B	Std. Error	Lower Bound	Upper bound	Tolerance	VIF
(constant)	0.182	0.157	-0.190	0.554		
x1	-2.002	2.963	-9.009	5.005	0.000	5868.654
x2	3.461	2.957	-3.530	10.452	0.000	5868.654
a. Dependent Variable : y2						
Model	Unstandardized Coefficients 95% Confidence interval for B Collinearity Statistic					
	B	Std. Error	Lower Bound	Upper bound	Tolerance	VIF
(constant)	-0.298	0.244	-0.875	0.279		
x1	1.526	4.596	-9.343	12.395	0.000	5868.654
x2	0.061	4.586	-10.783	10.905	0.000	5868.654
a. Dependent Variable : y3						

伍、多元共線性的症候

一般在迴歸分析或 SEM 分析中，多元共線性最主要的症狀為：不穩定或不可靠的迴歸係數估計值、偏高的變異數或共變數估計值、整體 R^2 具有顯著效果，但個別預測變項卻無顯著效果、負的變異量估計值、標準化的結構係數 >1 或 <-1（潛在變項與潛在變項間之相關）、偏高的未標準化迴歸係數之標準誤、過高的 MI（modification index）值、CI（condition index）值（>30）、雙變項間之積差相關 >.85 以上、多變項間之 VIF 過高（如 >10，亦可查看 Tolerance 值）。至於在 HLM 中，欲察覺多元共線性，可先檢查 HLM 報表中的 Tau 係數共變數與相關矩陣，看看非對角線相關係數是否接近 1 或 -1 或檢查共變數矩陣對角線元素是否接近 0，等於 0 代表相關斜率沒有隨機變異量。在 HLM 中，中心化可以移除隨機截距 & 斜率間

的高相關，以降低多元共線性，而避免 HLM 參數估計發生困難。

陸、多元共線性的發生原因與檢測方法

實務上，多元共線性的最常起因於：抽出樣本點侷限於一處時、變項數多於人數時、包含一個從其它預測變項衍生而來的變項（如社經地位係由教育程度與年收入計算而得）、不適當的類別變項編碼，如使用兩個虛擬變項代表性別（重複資訊）、使用相同或幾乎相同的預測變項（如個人收入與全家收入）、預測變項間具有線性關係（雙變項或多變項間）、交互作用項與高次方項等等。另外，Wissmann, Toutenburg, & Shalabh（2007）與 Liao（2010）的實徵研究亦發現，正確的選擇類別變項當參照組，也可降低多元共線性。假如研究者選擇反映次數較低的類別變項當參照組（low frequency as the reference），將會使得該類別變項與截距產生共線性。因此，在虛擬變項的迴歸分析中，選擇反映次數較低的類別變項當參照組就容易產生多元共線性。

以下將以最簡單的雙變數 X_1 & X_2 之迴歸分析模式： $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ 為例，推衍及說明多元共線性發生的原因與現象。本迴歸方程式中迴歸係數之變異數計算（O'Brien, 2007），如公式五所示。

$$\begin{aligned} & \hat{\beta}_j \text{ 的變異數} \\ \text{var}(\hat{\beta}_j) &= \sigma^2 \left(\frac{1}{1-r_{12}^2} \right) \frac{1}{S_{x_j}}, j=1,2 \quad (\text{公式五}) \\ \text{式中 } S_{x_1} &= \sum (x_{1i} - \bar{x}_1)^2, S_{x_2} = \sum (x_{2i} - \bar{x}_2)^2 \end{aligned}$$

當公式五中 $r_{12}=0$ ， β_j 的變異數會極小化，當 $r_{12} \cong 1$ ， β_j 的變異數會高度虛胖。例如當 $r_{12}=.95$ 時， β_j 的變異數（與 $r_{12}=0$ 相比）會增為 20 倍。由公式五可推知，假如（1）誤差項變異誤（ σ^2 ）很大，（2）離均差平方和不大（反映變異量，例如 S_{x_1} ），（3）兩變項的相關很高， β_j 的變異數就會偏高。因此，即使兩變項的相關很高，但如誤差項變異量很小而離均差平方和很大時，也不會有標準誤過高的問題。此外，因為誤差項變異誤（ $\sigma^2 = \frac{(1-R_{full}^2) \sum (Y - \bar{Y})^2}{N-k-1} = \frac{SS_{res}}{N-k-1}$ ）涉及決定係數（ R^2 ）、離均差與樣本大小（ N ），所以當 N 、離均差平方和不大、 R^2 愈大時，Beta 係數的估計值也會愈穩定。

如遇預測變項超過兩個以上時，公式五中的 $1-r_{12}^2$ 可以延伸為 $1-R_j^2$ （容忍度，tolerance）， R_j^2 係以預測變項 X_j 為效標，其餘 $p-1$ 個變項當作預測變項之決定係數。

如以相關係數呈現時（先將 X 變項中心化後，再將每一 X 變項向量化為單位長度，亦即加以正規化）， $y = \beta_1 x_1^* + \beta_2 x_2^* + \varepsilon$ 迴歸係數之解，可經由公式六~八分項逐步求解，帶入公式六。

$$XX = \begin{pmatrix} 1 & r_{12} \\ r_{12} & 1 \end{pmatrix} \quad (\text{公式六})$$

$$(XX)^{-1} = \begin{pmatrix} \frac{1}{1-r_{12}^2} & \frac{-r_{12}}{1-r_{12}^2} \\ \frac{-r_{12}}{1-r_{12}^2} & \frac{1}{1-r_{12}^2} \end{pmatrix} \quad (\text{公式七})$$

$$Xy = \begin{pmatrix} r_{1y} \\ r_{2y} \end{pmatrix} \quad (\text{公式八})$$

由公式七與八可以估計公式九之 β 值。

$$\hat{\beta} = (XX)^{-1}Xy = \begin{bmatrix} \frac{1}{1-r_{12}^2}(r_{1y} - r_{12}r_{2y}) \\ \frac{1}{1-r_{12}^2}(r_{2y} - r_{12}r_{1y}) \end{bmatrix} \quad (\text{公式九})$$

Beta變異數共變數矩陣：

$$\text{var}(\hat{\beta}) = \sigma^2(XX)^{-1} = \sigma^2C = \sigma^2 \begin{pmatrix} \frac{1}{1-r_{12}^2} & \frac{-r_{12}}{1-r_{12}^2} \\ \frac{-r_{12}}{1-r_{12}^2} & \frac{1}{1-r_{12}^2} \end{pmatrix} \quad (\text{公式十})$$

由公式十 & 十一可知， β_j 的變異數的計算涉及兩大要素：估計誤差項變異誤（ σ^2 ）與 $(XX)^{-1}$ 。因此，假如 $(XX)^{-1}$ 的值逼近於0時，將使得 β_j 的變異數估計值非常不穩定。Gatignon & Vosgerau(2005), Echambadi & Hess(2007)研究證實 $(XX)^{-1}$ 不會因資料是否置中而改變其值，因而影響 β_j 的變異數估計值的另一因素：估計誤差項變異誤（ σ^2 ），亦不可等閒視之（尤其在今日電腦的高度精確計算能力之下， $(XX)^{-1}$ 的干擾已可以藉著運算精確度之提高顯著減輕）。

$$\text{當 } r_{12} \rightarrow 1 \quad \text{var}(\hat{\beta}_j) = \sigma^2 c_{jj} \rightarrow \infty, \text{ 而 } \text{cov}(\hat{\beta}_1, \hat{\beta}_2) = \sigma^2 c_{12} \rightarrow \infty. \quad (\text{公式十一})$$

$$VIF_j = c_{jj} = \frac{1}{1-R_j^2} \quad j = 1, 2, 3 \dots p \quad (\text{公式十二})$$

公式十二係 VIF 之計算公式，式中 R_j^2 係以預測變項 X_j 為效標，其餘 $p-1$ 個

變項當作預測變項之決定係數。由公式十一&十二可知， β_j 的變異數之計算涉及估計誤差項變異誤 (σ^2) 與 VIF 加以計算；因而研究者假如僅根據 VIF，而忽視估計誤差項變異誤，去論斷多元共線性的問題，有可能失之偏頗。

綜上所述，多元共線性之診斷需進行全面性之分析，包含以下幾個重要因素：樣本大小、預測變項之變異量、變項間之相關性、整體變項之預測力與估計誤差變異量，這些因素均會造成迴歸係數估計標準誤過大而導致參數估計值不穩定現象。

過去許多論文文獻僅論及中心化可以降低線性變項與交互作用項間之共變量，以改善共線性問題；但對於原始預測變項間之共線性，如果建立交互作用項的方法錯誤，中心化根本無法降低原始預測變項間多元共線性之現象。請看以下圖 4 之實例說明：X3 係 X1 & X2 之交乘積項，X6 係 X3 的離均差分數的變項 ($\bar{X}_3=11$)。

	x1	x2	x3	x5	x6	y
1	3	4	12	.40	1.00	12.00
2	4	3	12	-.60	1.00	4.00
3	3	3	9	-.60	-2.00	18.00
4	2	2	4	-1.60	-7.00	14.00
5	3	6	18	2.40	7.00	21.00
6						

圖 4 原始預測變項間之共線性範例

圖 4 範例中，X5 & x6 係 x2 & x3 之中心化新變項，其相關矩陣如表 3 所示。

表 3 中心化與未中心化相關矩陣

		y	x2	x3
Pearson Correlation	y	1.000	0.497	0.219
	x2	0.497	1.000	0.938
	x3	0.219	0.938	1.000
		y	x5	x6
Pearson Correlation	y	1.000	0.497	0.219
	x5	0.497	1.000	0.938
	x6	0.219	0.938	1.000

因為置中後變項內之數據之相對位置並未改變，變項間之相關係數並不因置中與否而改變。讀者可以發現 X2 & X3 相關高達.938，而 X2 & X3 置中後之 X5 & X6 相關仍為.938（注意置中變項因為相對位置未變，其相關係數不會改變），相關之迴歸係數估計值與 VIF 均為 8.263，置中之前後均未改變。由上述公式十二可推知，當雙變項間之相關係數未變，其相關之 VIF 也不會改變（參見表 4 右側欄位內容）。

前述表 3 中變項間之相關係數並不因置中與否而改變，可以利用公式十三~公式十五之推演而證實。如就置中前 Pearson 積差相關的不偏估計值公式來看，其基本公式為

$$r_{\text{前}} = \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{2i} - \bar{X}_2)(X_{3i} - \bar{X}_3)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{2i} - \bar{X}_2)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{3i} - \bar{X}_3)^2}} \quad (\text{公式十三})$$

置中後 Pearson 積差相關的公式則如公式十四所示：

$$\begin{aligned} r_{\text{後}} &= \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{2i}^c - \bar{X}_2^c)(X_{3i}^c - \bar{X}_3^c)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{2i}^c - \bar{X}_2^c)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{3i}^c - \bar{X}_3^c)^2}} \\ &= \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{2i}^c - 0)(X_{3i}^c - 0)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{2i}^c - 0)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{3i}^c - 0)^2}} \\ &= \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{2i}^c X_{3i}^c)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{2i}^c)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{3i}^c)^2}} \\ &= \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{2i} - \bar{X}_2)(X_{3i} - \bar{X}_3)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{2i} - \bar{X}_2)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{3i} - \bar{X}_3)^2}} = r_{\text{前}} \end{aligned} \quad (\text{公式十四})$$

公式十四中，

$$\begin{aligned} X_{2i}^c &= (X_{2i} - \bar{X}_2), \bar{X}_2^c = \frac{\sum_{i=1}^n (X_{2i} - \bar{X}_2)}{n} = \frac{\sum_{i=1}^n X_{2i} - n\bar{X}_2}{n} = 0 \\ X_{3i}^c &= (X_{3i} - \bar{X}_3), \bar{X}_3^c = \frac{\sum_{i=1}^n (X_{3i} - \bar{X}_3)}{n} = \frac{\sum_{i=1}^n X_{3i} - n\bar{X}_3}{n} = 0 \end{aligned} \quad (\text{公式十五})$$

上述數學公式之推演，說明了交互作用項的建置方法如採交乘積項之離差分

數，並無法改變原始變項與乘積項間之關係。如欲改善原始變項與乘積項間之關係，可能必須使用原始變項的離差分數之交乘積作為交乘積項（參閱公式二十一）。

表 4 中心化與未中心化迴歸分析結果

Model	Unstandardized Coefficients				Collinearity Statistics	
	B	Std. Error	t	Sig.	Tolerance	VIF
(Constant)	5.225	5.806	0.900	0.463		
X2	10.340	4.325	2.391	0.139	0.121	8.263
X3	-2.604	1.286	-2.025	0.180	0.121	8.263
Model	Unstandardized Coefficients				Collinearity Statistics	
	B	Std. Error	t	Sig.	Tolerance	VIF
(Constant)	13.800	2.041	6.761	0.021		
X5	10.340	4.325	2.391	0.139	0.121	8.263
X6	-2.604	1.286	-2.025	0.180	0.121	8.263

a. Dependant Variable : y

另由公式十可知，假如 $(XX)^{-1}$ 的值逼近於 0，將使得 β_j 的變異數估計值非常不穩定。茲以圖 5 之例子說明之，例子中 x1 & x2 之相關係數逼近於 0，導致 VIF 值逼近 1。照道理一般均直覺會認為不會有多元共線性問題；但此迴歸分析卻有嚴重的多元共線性現象。

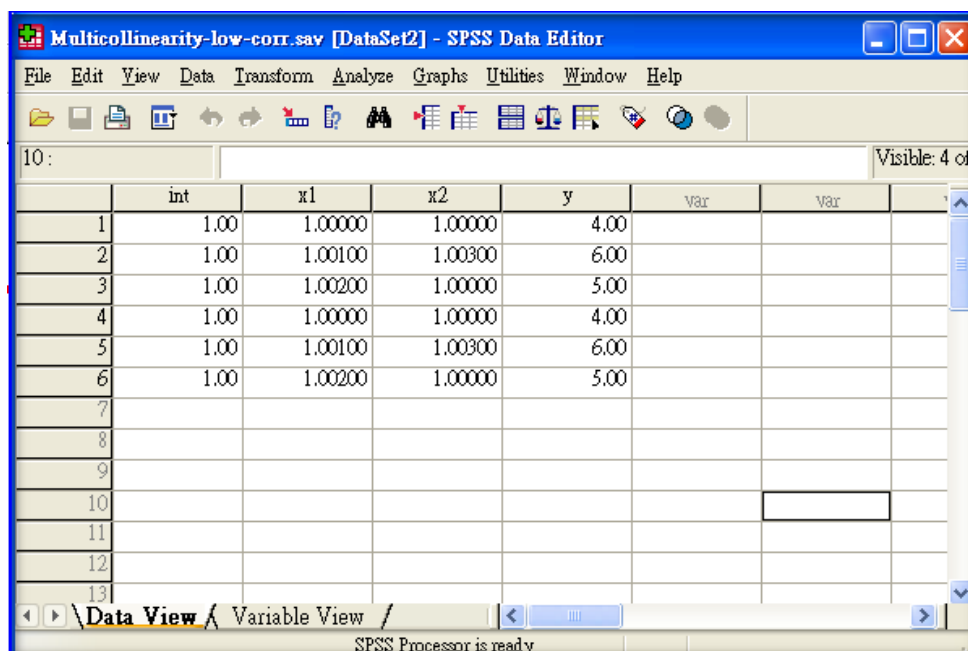


圖 5 相關係數逼近於 0 的多元共線性實例

從表 5 右上角之迴歸分析之共線性統計量 VIF=1 來看，本迴歸分析沒有多元

共線性問題；但如從 $(XX)^{-1}$ 的值逼近於 0 ($2.879999998 \times \frac{1}{10^{-10}}$) 來看，本筆資料卻有嚴重多元共線性問題；又如從高得離譜之 CI 指標 (1426.688 & 3160.470) 來看，也反映出本迴歸分析有嚴重的多元共線性問題，不同的指標出現了矛盾的統計診斷結果。因此，低的相關、低的 VIF 並不保證不會有多元共線性的症狀，CI、變異比均可作為輔助指標 (Chennamaneni, Echambadi, Hess, et al., 2008)。由公式五

中的第三個要素 $\frac{1}{S_{x_j}}$ 可以推知，如果預測變項的離散量過小 (如本例 x1 & x2 的標準差分別為 .00089443 & .00154919) 就會間接導致 CI 指標過高。反之，高的 VIF 只是多元共線性的充分條件而非必要條件 (Belsley, 1991)。由此觀之，VIF 指標旨在反映預測變項間之相關程度，而 CI、變異比率指標則在反映預測變項的離散量。研究者必須使用多元之診斷指標，才能較正確診斷出多元共線性的問題是否存在 (Echambadi, Hess, 2007)。

表 5 相關係數逼近於 0 的多元共線性實例之迴歸分析結果

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-996.000	.000		-5091442	.000		
	x1	500.000	0.0001692446573317	.500	2954303.0	.000	1.000	1.000
	x2	500.000	.000	.866	5117002.9	.000	1.000	1.000
a. Dependent Variable: y								

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions		
				(Constant)	x1	x2
1	1	3.000	1.000	.00	.00	.00
	2	1.47E-006	1426.688	.03	.12	.88
	3	3.00E-007	3160.470	.97	.88	.12
a. Dependent Variable: y						

另外，一般都會誤認為雙變項的相關不高，就不會發生多元共線性；事實上，這是一種迷思。茲以圖 6 之資料為例釐清此迷思，該資料含有四個預測變項 x1~x4。這四個預測變項 x1~x4 間之相關矩陣如表 6 所示。

	Y	x1	x2	x3	x4	tot
1	6	8	1	1	1	10.00
2	5	8	1	1	0	10.00
3	3	8	1	1	0	10.00
4	2	0	0	9	1	9.00
5	3	0	0	9	1	9.00
6	1	0	0	9	1	9.00
7	3	2	7	0	1	9.00
8	4	2	7	0	1	9.00
9	5	2	7	0	1	9.00
10	2	0	0	0	10	.00
11	1	0	0	0	10	.00
12	2	0	0	0	10	.00

圖 6 雙變項相關係數不高的多元共線性實例

表 6 預測變項 x1~x4 間之相關矩陣

	x1	x2	x3	x4	Tot*
x1	1.000	0.052	-0.343	-0.498	-.997
x2	0.520	1.000	-0.432	-0.371	-.996
x3	-0.343	-0.432	1.000	-0.355	-.997
x4	-0.498	-0.371	-0.355	1.000	-.998

*註：Tot 係校正後之題目與總分之相關；例如：以 X1 為例，Tot=x2+x3+x4

由表 6 之四個預測變項 x1~x4 間之相關矩陣結果知，這四個預測變項之相關介於-.498~.520，雙變項間之相關係數並不高。但此迴歸分析卻出現嚴重之多元共線性現象，根本無法經由檢視相關矩陣加以判定。

因為整體來看，本模式具有顯著之預測效果（ $R^2=.769$ ， $p=.022$ ），但個別預測變項卻無顯著之預測效果（ $p>.05$ 以上），各預測效果之 VIF 值均非常大（ >158 ），詳見表 7 之內容。

本例會發生多元共線性現象的主因，係以其它預測變項預測第 i 個預測變項之決定係數過高所致（例如： $x_4=10-.957x_1-1.012x_2-1x_3$ ， $R^2=.997$ ），亦即容忍度（tolerance）過小所致。

換言之，任何預測變項不可与其它預測變項之組合具有高度相關（如本例中 x_4 與 tot 具有高度相關：-.998），否則也會發生多元共線性現象。當一個預測變項与其它預測變項之組合具有高度相關時，研究者必須靠統計指標（如 VIF/Tolerance）加以檢驗，並無法直接從相關矩陣中辨識出來。

表 7 預測變項 x1~x4 間之迴歸分析結果

Model	Unstandardized Coefficients		Standardized Coefficients		Collinearity Statistics		
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
(Constant)	-18.333	11.966		-1.532	0.169		
x1	2.254	1.147	4.762	1.965	0.090	0.006	178.287
x2	2.261	1.215	4.246	1.861	0.105	0.006	158.046
x3	2.037	1.199	4.954	1.700	0.133	0.004	257.907
x4	2.000	1.195	5.166	1.673	0.138	0.003	289.375
Model Dimension Eigenvalues	Condition Index	Variance Proportions					
		(Constant)	x1	x2	x3	x4	
1	2.383	1.000	0.00	0.00	0.00	0.00	0.00
2	1.098	1.473	0.00	0.00	0.00	0.00	0.00
3	0.914	1.615	0.00	0.00	0.00	0.00	0.00
4	0.605	1.984	0.00	0.00	0.00	0.00	0.00
5	0.000	86.580	1.00	1.00	1.00	1.00	1.00

a. Dependent Variable : y

表 7 中 CI (Condition Index) 指標係利用 $(X'X)$ 之交乘積矩陣特徵值 (eigenvalues) 計算而來，當特徵值愈小 (如接近 0) 即反映出線性相依程度愈嚴

重。CI 指標的計算公式：
$$= \sqrt{\frac{\lambda_{\max}}{\lambda_i}}$$
，當 CI 值大於 30 時 (本例有一特徵值逼近於 0，其 CI 值為 86.58)，即表示該筆資料具有嚴重的多元共線性。表 7 中變異比率 (Variance proportion) 指標，係表中每一特徵值的特徵向量，每一特徵向量反映該變項對於相關特徵值的貢獻量，每一縱行之變異比和為 1，當一個特徵向量相關之變異比中，有兩個以上變項的變異比超過 0.5 以上，即表示存在多元共線性 (Judge, et al., 1985)；例如，在表 7 底部向度 5 中，該特徵值為 0.000 且在 x1~x4 上之變異比率均為 1，即表這四個變項間會產生嚴重的多元共線性問題。

柒、為何變項置中並不一定會減輕多元共線性的威脅？

Belsley (1984), Gatignon & Vosgerau (2005), Echambadi & Hess (2007) 均認為交互作用項的平均數置中，雖有時可降低單一變項與交互作用項間之線性相關，但並無法解決多元共線性問題。這些研究證實了中心化變項既無法改善參數估計值之運算精確性，亦無法提高整體 R^2 之解釋力 (即使 VIF 下降)，因為交乘積矩陣 $(X'X)$ 的行列式值在中心化與未中心化的資料上均完全相同。

前述之論述似乎似是而非，筆者乃根據他們的三種論述證據，逐一加以公式推導，並輔以實例 (參見表 8~表 9 與圖 8~圖 10) 驗證之。

證據一：行列式完全相同

公式十六係一原始變項間之交乘積矩陣，含有截距、變項 X1、變項 X2 及交互作用項 X12。T 表示總人數。

$$|A| = |XX| = \begin{bmatrix} T & T\bar{X}_1 & T\bar{X}_2 & T\bar{X}_{12} \\ T\bar{X}_1 & X'_1X_1 & X'_1X_2 & X'_1X_{12} \\ T\bar{X}_2 & X'_2X_1 & X'_2X_2 & X'_2X_{12} \\ T\bar{X}_{12} & X'_{12}X_1 & X'_{12}X_2 & X'_{12}X_{12} \end{bmatrix} \quad (\text{公式十六})$$

上述公式十六的矩陣如果經過以下行列式的基本運算：將第一列（1st row）乘上 $-\bar{X}_1$ 加到第二列、第一列乘上 $-\bar{X}_2$ 加到第三列、第一列乘上 $-\bar{X}_1\bar{X}_2$ 加到第四列；再將第一行（1st column）乘上 $-\bar{X}_1$ 加到第二行、第一行乘上 $-\bar{X}_2$ 加到第三行、第一行乘上 $-\bar{X}_1\bar{X}_2$ 加到第四行，最後將可得到：

$$|XX| = \begin{bmatrix} T & 0 & 0 & T(\bar{X}_{12} - \bar{X}_1\bar{X}_2) \\ 0 & (X'_1X_1 - T\bar{X}_1^2) & (X'_1X_2 - T\bar{X}_1\bar{X}_2) & (X'_{12}X_1 - T\bar{X}_1\bar{X}_{12}) \\ 0 & (X'_2X_1 - T\bar{X}_1\bar{X}_2) & (X'_2X_2 - T\bar{X}_2^2) & (X'_{12}X_2 - T\bar{X}_2\bar{X}_{12}) \\ T(\bar{X}_{12} - \bar{X}_1\bar{X}_2) & (X'_{12}X_1 - T\bar{X}_1\bar{X}_{12}) & (X'_{12}X_2 - T\bar{X}_2\bar{X}_{12}) & (X'_{12}X_{12} - 2T\bar{X}_1\bar{X}_2\bar{X}_{12} + T\bar{X}_1^2\bar{X}_2^2) \end{bmatrix} \quad (\text{公式十七})$$

接著，繼續將上述公式十七矩陣之第二列乘上 $-\bar{X}_2$ 、第三列乘上 $-\bar{X}_1$ ，分別加到第四列；再將第二行乘上 $-\bar{X}_2$ 、第三行乘上 $-\bar{X}_1$ ，分別加到第四行（其中， $\bar{X}_1X'_2X_{12} = \bar{X}_1T\bar{X}_{12}$ ， $\bar{X}_2X'_1X_{12} = \bar{X}_2T\bar{X}_{12}$ ）；最後可得：

$$|XX| = \begin{bmatrix} T & 0 & 0 & T(\bar{X}_{12} - \bar{X}_1\bar{X}_2) \\ 0 & (X'_1X_1 - T\bar{X}_1^2) & (X'_1X_2 - T\bar{X}_1\bar{X}_2) & \begin{pmatrix} X'_{12}X_1 - \bar{X}_2X'_1X_1 \\ + 2T\bar{X}_1^2\bar{X}_2 - 2T\bar{X}_1\bar{X}_{12} \end{pmatrix} \\ 0 & (X'_2X_1 - T\bar{X}_1\bar{X}_2) & (X'_2X_2 - T\bar{X}_2^2) & \begin{pmatrix} X'_{12}X_2 - \bar{X}_1X'_2X_2 \\ + 2T\bar{X}_1\bar{X}_2^2 - 2T\bar{X}_2\bar{X}_{12} \end{pmatrix} \\ T(\bar{X}_{12} - \bar{X}_1\bar{X}_2) & \begin{pmatrix} X'_{12}X_1 - \bar{X}_2X'_1X_1 \\ + 2T\bar{X}_1^2\bar{X}_2 - 2T\bar{X}_1\bar{X}_{12} \end{pmatrix} & \begin{pmatrix} X'_{12}X_2 - \bar{X}_1X'_2X_2 \\ + 2T\bar{X}_1\bar{X}_2^2 - 2T\bar{X}_2\bar{X}_{12} \end{pmatrix} & \begin{pmatrix} X'_{12}X_{12} + \bar{X}_2^2X'_1X_1 + \bar{X}_1^2X'_2X_2 \\ - 2\bar{X}_1X'_{12}X_2 - 2\bar{X}_2X'_1X_{12} \\ + 4T\bar{X}_1\bar{X}_2\bar{X}_{12} - 3T\bar{X}_1^2\bar{X}_2^2 \end{pmatrix} \end{bmatrix}$$

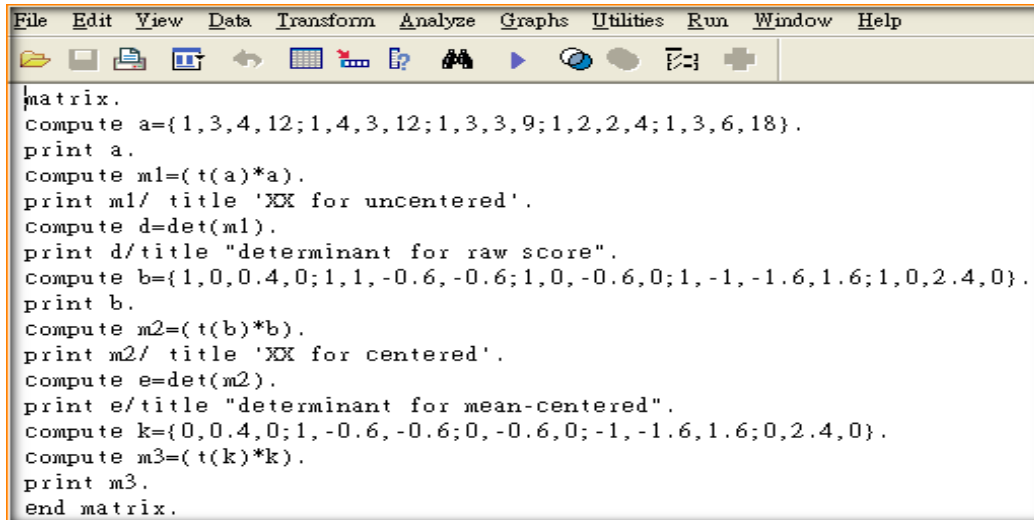
$$= |X^d X^d| = |B| \quad (\text{公式十八})$$

公式十八中，

$$X^d = X - \bar{X} \quad (\text{公式十九})$$

若 B 矩陣（假設為置中矩陣）是 A 矩陣（假設為未置中矩陣）的某一行（或行）的倍數加到 A 矩陣另一行（或行），則兩個矩陣的行列式值將相等（ $|A| = |B|$ ）。

另外，讀者不難發現離差分數的交乘積矩陣 $|X^d X^d|$ ，可由原始變項的原始分數直接計算出來。據此，可以導出前述之 $X'X$ 矩陣未中心化與中心化後之行列式值完全相同；Gatignon & Vosgerau (2005) 與 Echambadi & Hess (2007) 的模擬研究結果也證實這項結果。請檢閱圖 7 中之 A 矩陣實例，A 含有截距、X1 & X2 變項與交互作用項 X1X2。再利用圖 7 之 SPSS 矩陣語言，就可跑出該 $X'X$ 矩陣未中心化與中心化之行列式值，而知兩者之行列式值均為 4。因此，也體現了置中對多元共線性沒有任何效用，但也無任何傷害。



```
matrix.
compute a={1,3,4,12;1,4,3,12;1,3,3,9;1,2,2,4;1,3,6,18}.
print a.
compute m1=(t(a)*a).
print m1/ title 'XX for uncentered'.
compute d=det(m1).
print d/ title "determinant for raw score".
compute b={1,0,0.4,0;1,1,-0.6,-0.6;1,0,-0.6,0;1,-1,-1.6,1.6;1,0,2.4,0}.
print b.
compute m2=(t(b)*b).
print m2/ title 'XX for centered'.
compute e=det(m2).
print e/ title "determinant for mean-centered".
compute k={0,0.4,0;1,-0.6,-0.6;0,-0.6,0;-1,-1.6,1.6;0,2.4,0}.
compute m3=(t(k)*k).
print m3.
end matrix.
```

圖 7 SPSS 矩陣語言：計算置中前後 $X'X$ 的行列式值

不過，Shieh (2011) 認為含有截距項的行列式值 (the augmented data product determinant) 並不能用來評估變項置中對於多元共線性的影響程度；仍須參考其它指標或檢驗置中後交互作用項與原始變項間之相關係數。他舉了兩個實際例子 (參見該文章之表 2 & 表 4 資料) 說明：置中後交互作用項與原始變項間之相關係數一定會受到影響，可能變大也可能變小 (雖然其含截距項交乘積項之行列式值均相同)，請看以下圖 8 & 圖 9 之實例說明。

證據二：置中後交互作用項與原始變項間相關係數不一定變小。

公式二十係雙變數 (原始變項 X1 與交互作用項 X1X2) 未置中前的 Pearson 積差相關係數公式：

$$r_{前} = \frac{\text{雙變數之共變數}}{\text{雙變數之標準差}} = \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{1i}X_{2i} - \bar{X}_{1i}X_{2i})}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{1i} - \bar{X}_1)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{1i}X_{2i} - \bar{X}_{1i}X_{2i})^2}} \quad (\text{公式二十})$$

$$X_{1i}^c = (X_{1i} - \bar{X}_1), \bar{X}_1^c = \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)}{n} = \frac{\sum_{i=1}^n X_{1i} - n\bar{X}_1}{n} = 0$$

$$X_{3i}^c = (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2), \bar{X}_3^c = \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{n} \neq 0 \quad (\text{公式二十一})$$

請注意公式二十一中交互作用項之建立，係來自 X1 與 X2 的離差分數之交乘積，而非 X1 與 X2 交乘積項之離差分數（請參照比較公式十五）。公式二十二係雙變數置中後的 Pearson 積差相關係數公式：

$$r_{\text{後}} = \frac{\text{雙變數之共變數}}{\text{雙變數之標準差}} = \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{1i}^c - \bar{X}_1^c)(X_{3i}^c - \bar{X}_3^c)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{1i}^c - \bar{X}_1^c)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{3i}^c - \bar{X}_3^c)^2}}$$

$$= \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{1i}^c - 0) \left((X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) - \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{n} \right)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{1i}^c - 0)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left((X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) - \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{n} \right)^2}}$$

$$= \frac{\frac{1}{n-1} \sum_{i=1}^n (X_{1i} - \bar{X}_1) \left((X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) - \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{n} \right)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{1i} - \bar{X}_1)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left((X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) - \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{n} \right)^2}}$$

$$\neq r_{\text{前}} \quad (\text{公式二十二})$$

經仔細推敲公式二十 & 公式二十二的預測項之平均數與變異量，發現 X1 變數與交互作用項之相關大小主要受制於「X1 變數與交互作用項」的數據離散特質。尤其當此交互作用項， $(X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)$ ，的標準差與平均數之比值（仿變異係數，或稱為相對差異係數的做法）為原始變項之交互作用項， $(X_{1i})(X_{2i})$ ，的標準差與平均數之比值差異遠大於 1 時，就會出現置中後一階變項與交互作用項間之相關上升現象。換言之，如變項置中後出現交互作用項之相對變異量異常增大（常出現在含有極端值時），就會出現置中後一階變項與交互作用項間之相關不降反升之現象。例如：圖 8 中 x1、x2 係原始變項，X3 為 x1、x2 之交乘積項；x4~x6 係置中後之相對應變項；y 係效標變項。

	x1	x2	x3	x4	x5	x6	y
1	3	4	12	0	.40	.00	12.00
2	4	3	12	1	-.60	-.60	4.00
3	3	3	9	0	-.60	.00	18.00
4	2	2	4	-1	-1.60	1.60	14.00
5	3	6	18	0	2.40	.00	21.00
6	6	6	36	0	3.60	.00	42.00

圖 8 置中前後相關係數不一定變小實例

圖 8 中未置中交互作用變項的平均數與標準差分別為 11 與 5.099，而置中交互作用變項的平均數與標準差分別為：0.2 與 0.82462。據此，未置中交互作用變項的變異比（又稱為相對差異係數，coefficient of relative variability）為：5.099/11=0.4635；而置中後交互作用變項的變異比變為：0.8242/0.2=4.121；此變異比的比較，最適合於比較兩群體的平均數具有重大差異時。由此可知：置中前後交互作用項本身的「變異比差異」異常大於 1（相對變異比超過 8 倍之多）。因為變異比相對的異常增大，因此可以斷言：一階變項與交互作用項間之相關至少有一對會提高。表 8 係圖 8 中各變項之相關矩陣，注意置中變項間之相關係數不一定會下降。例如， $|r_{x1 \times 3}| = .555 < |r_{x4 \times 6}| = -.943$ ，亦即置中後一階變項與交互作用項間之相關不降反升，而無法達到降低多元共線性對於參數估計值穩定性之威脅。Shieh(2011)的模擬研究結果（參見該文章之表 4 資料，含有極端值）也發現，中心化變項會因資料結構之屬性，反而提高預測變項與交互作用項間之相關性及相關之 VIF 值；也與筆者的變異比理論相吻合。因此，研究者不可將中心化變項視為減輕多元共線性的萬靈丹，中心化是否能有效降低多元共線性，端視交互作用項的標準差與平均數之「變異比」是否變大或變小而定：當置中後其變異比如出現增大，中心化變項將無法有效降低原始變項與交互作用項間之多元共線性現象。

表 8 置中前後相關係數矩陣：相關係數會變大

		y	X1	X2	X3
Pearson Correlation	y	1.000	-0.544	0.497	0.219
	x1	-0.544	1.000	0.233	0.555
	x2	0.497	0.233	1.000	0.938
	x3	0.219	0.555	0.938	1.000
		y	x4	x5	x6
Pearson Correlation	y	1.000	-0.544	0.497	0.289
	x4	-0.544	1.000	0.233	-0.943
	x5	0.497	0.233	1.000	-0.440
	x6	0.289	-0.943	-0.440	1.000

為了進一步檢驗「變異比差異」立論的正確性，利用 Shieh (2011) 文中的表 2 資料（參見圖 9 內的數據），將置中前後的相關矩陣列述於表 10。圖 9 中之未置

中交互作用變項的平均數與標準差分別為：8.478 與 27.206，而置中交互作用變項的平均數與標準差分別為：0.369 與 0.341。據此，未置中交互作用變項的變異比為： $27.206/8.478=3.21$ ；而置中後交互作用變項的變異比變為： $0.341/0.369=.924$ 。因置中後交互作用項本身的變異比值（.924）並未異常變大，再細查該筆資料亦未發現有極端值出現。因此，可以預測一階變項與交互作用項間之相關會下降；中心化變項將可以有效降低原始變項與交互作用項間之多元共線性現象。

	X	Z	XZ	XC	ZC	XZC
1	6.800	5.600	38.080	1.450	.570	.827
2	5.000	5.100	25.050	-.350	.070	-.025
3	4.700	3.700	17.390	-.650	-1.330	.865
4	5.000	4.900	24.500	-.350	-.130	.046
5	4.700	4.700	22.090	-.650	-.330	.215
6	5.100	3.700	18.870	-.250	-1.330	.333
7	5.800	6.800	39.440	.450	1.770	.797
8	3.600	5.100	18.360	-1.750	.070	-.123
9	6.800	5.100	34.680	1.450	.070	.102
10	6.000	5.600	33.600	.650	.570	.371

圖 9 置中前後相關係數會變小實例：取自 Shieh (2011)，表 2 資料

表 9 置中前後相關係數矩陣：相關係數會變小

	X	Z	XZ	XC	ZC	XZC
X	1	.414	.853**	1.000**	.414	.410
Z	.414	1	.827**	.414	1.000**	.139
XZ	.853**	.827**	1	.853**	.827**	.373
XC	1.000**	.414	.853**	1	.414	.410
ZC	.414	1.000**	.827**	.414	1	.139
XZC	.410	.139	.373	.410	.139	1

由表 9 知，置中後一階變項與交互作用項間之相關已下降（.853→.410，.827→.139），而達到降低多元共線性對於參數估計值穩定性之威脅。為降低前述之結果係因機遇所致（capitalized by chance），筆者乃利用表 10 之 SPSS 語法程式，採用 SPSS 內建函數 RV.Normal（平均數，標準差），模擬雙變數之資料點，再利用 SPSS 之統計分析功能，計算各變項置中前後預測變數與交互作用項之平均數與標準差的比值，並計算各變項置中前後之相關矩陣。經實徵檢驗數

個來自不同抽樣分配的資料集，亦發現所得結果與上述之理論相吻合：置中前與置中後之原始預測變項或其交互作用項的標準差與平均數之變異比差異，因係反映出該變項數據置中前後的相對離散量，可以預測置中是否可以有效降低多元共線性的威脅。因此，筆者所提出的預測變項「變異比理論」獲得實徵性的驗證。

表 10 SPSS 語法程式：模擬雙變數之資料點

```
comment 'A SPSS program for RELATIVE VARIANCE RATIO experiment'.
comment 'Written by Fred Li, 2012'.
comment 'Data is generated from NORMAL distribution WITH specified mean & sd'.
Input program.
DATA LIST /#X1 TO #X5 1-5 #Y1 TO #Y5 6-10.
Loop #I=1 to 5.
VECTOR A=#X1 TO #X5.
COMPUTE X=A (#I) .
VECTOR B=#Y1 TO #Y5.
COMPUTE Y=B (#I) .
End Case.
End Loop.
End Input Program.
BEGIN DATA
1
END DATA.
Set Seed Random.
LOOP #I=1 TO 5.
Compute X=rv.normal (1,5) .
Compute Y=rv.normal (1,5) .
End Loop.
Print/X Y 1-4 (2) .
EXE.
```

置中前與置中後之原始預測變項與其交互作用項的標準差與平均數之變異比差異，反映出該變項數據置中前後的相對離散量。通常，離散量愈大，雙變數的相關愈不會出現全距減縮（range-restriction）現象。一般來說，置中變項的原始預測變項與交乘積變項應會出現全距減縮現象，以致於使得相關係數下降。因此，假如原始預測變項與交互作用項的標準差與平均數之變異比變得異常大，就很有可能使得相關係數上升。前述之「變異比差異」理論，經理論之論述與模擬實例之驗證，可以詮釋變項置中可能降低也可能提高一階變項與交互作用項間相關之現象。

證據三：迴歸模式置中前、後係等同統計模式（相同 R²、相同適配度值），參數可以互換。

Echambadi, Hess (2007) 的論文中，曾以表 1 的消費評鑑例子（p.442，此例子出自 Bottomley & Holden, 2006），進行兩個模式參數可以互換的說明。本例迴歸分析旨在評鑑消費者對於品牌替代性（substitutability, x1）與品質（quality, x2）在購置品牌替代品與其所覺知品質上的預測力。表 11 左右兩側係置中前後該迴歸模

式參數估計值之對照表。表 11 上半部第二行示範，置中後 Substitute 參數($b1=0.132$)可以利用表 11 左側未置中前之相關參數計算而得 ($b1 = a_1 + \bar{x}_2 a_3$ ，推衍細節參見該文之附錄)；其相關之標準誤 ($SEb1=0.008$ ，參見表 11 底部)。整體之 $R^2 (=0.30)$ 在這兩個模式中均無不同，因而為等同統計模式。統計上，所謂等同統計模式乃指相同適配度與預測值，參數也可以互轉。根據未置中估計值來看，Substitute 參數 ($a1$) 係負值 (-0.04655)，意謂著高品牌替代性會導致較低的品牌替代品購買慾，這似乎違反常理。爰此，研究者可能會懷疑多元共線性的存在，尤其在檢視了置中估計值來看，Substitute 參數 ($a1$) 係出現正值 ($.132$)，符合了我們的預期。據此，研究者可能會相信置中已減輕了多元共線性的效應。不過，假如研究者仔細檢視未置中資料的各參數的標準誤並無顯著虛胖現象 ($.005 \sim .026$)。其實，這兩個模式中的參數代表著不同意義，是不能直接比較的。在未置中的迴歸模式裡，Substitute 參數的效果 ($a1$) 指的是當其它預測變項為 0 時的簡單效果 (simple effect)；在置中的迴歸模式裡，Substitute 參數的效果 ($b1$) 指的是當其它預測變項 (x_2) 為平均值時的主要效果 (main effect)；亦即產品品質中等時，即會有品牌替代品購買慾。因此，在未置中的迴歸模式裡，當品牌品質缺乏時，Substitute 參數的效果 ($a1$) 為負值是合理的，沒有品質的產品是無法行銷的。

表 11 置中前後仍係等同迴歸模式

$a_1 + \bar{x}_2 a_3 = -0.047 + 5.208 * 0.034 = 0.132$ Substitute 主要效果							
Uncentered estimate				Mean-centered estimate			
	Substitute a1	Quality a2	Substitute × Quality a3		Substitute b1	Quality b2	Substitute × Quality b3
Estimate	-0.04655	0.250	0.03426	Estimate	0.132	0.349	0.034
(SEj)	(0.026)	(0.015)	(0.005)	(SEj)	(0.008)	(0.009)	(0.005)
VIF	13.69	2.90	16.05	VIF	1.31	1.03	1.02
VAR-COVa				VAR-COVb			
a1	0.000656	-	-	b1	0.000063	-	-
a2	0.0003	0.00023	-	b2	-0.00001	0.000085	-
a3	-0.00011	-0.00005	0.000021	b3	-0.000003	0.000005	0.00002
Variable mean							
$\bar{x}_j$	2.898	5.2077		$\bar{x}_j$			
R2	0.30			R2	0.30		
$VAR_{a_1 + \bar{x}_2 a_3} = VAR_{a_1} + 2\bar{x}_2 COV_{a_1 a_3} + \bar{x}_2^2 VAR_{a_3}$ $= 0.000656 + 2 \times 5.2077 \times (-0.00011) + 5.2077^2 \times 0.000021 = 0.000062 = 0.0082$							

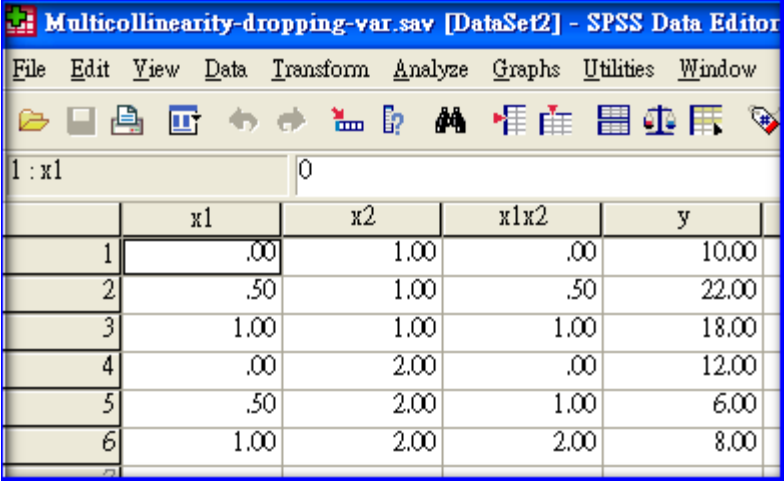
註：摘要自 Echambadi & Hess, 2007

綜上所述，變項置中並不一定會減輕多元共線性的威脅，此乃多元共線性之現象（不穩定的參數估計值或標準誤過大）與樣本大小、預測變項之變異量、變項間之相關性、行列式值、交互作用項之建立方法、整體變項之預測力具有密切關係；而中心化只會影響交互作用項或高次方項（當作預測變項）的變異量。中心化並不

會改變原始變項間之相關性與整體變項之預測力，因而對於迴歸係數估計值的變異量將毫無影響。至此，可獲得一結論：中心化會減輕多元共線性之最大原因，係預測變項或交互作用項或高次方項（當作預測變項）的「誤差變異量」，常因相關線性變項之中心化而顯著下降所致。此項結論亦可由公式十的兩個因素（估計誤差項變異量（ σ^2 ）與 $(XX)^{-1}$ ）獲得證實。

捌、多元共線性的解決方法

多元共線性是程度上而非全有或全無之問題，欲完全排除它乃是不可能的事。中心化並非解決多元共線性的唯一選擇，過去研究者（如 Berry & Feldman, 1985 & 1993; Yu, 1989）曾提出不少方案，例如：因為多元共線性是樣本問題（a sample problem），破解之道乃是蒐集更多、更佳之資料、刪除共線性變項、殘差中心化、進行問題預測變項之資料轉換、選擇高頻率的類別變項當作參照組、使用主成分迴歸分析、使用 Ridge 迴歸分析、合併高相關預測變項（假如具有理論基礎且變項間之相關不可過低）、中心化交互作用項或高次方項、使用較佳之研究設計及正確建立交互作用項。這些方法中，殘差中心化已被證實會偏估線性變項效果（Echambadi, et al., 2006），主成分迴歸分析因無適當統計方法可以考驗刪除後交互作用效果，亦不推薦使用它（Morris, Sherman, & Mansfield, 1986; Cronbach, 1987）；而最常用的刪除有問題線性預測變項，當變項間之相關過低時也不可行，可能會導致偏估之參數估計值與模式界定錯誤問題，參見圖 10 實例的進一步說明。Arceneaux & Huber（2007）亦曾主張因多元共線性導致估計值不精確，但仍比因刪除變項而導致偏差的估計值更勝一籌。難怪，Berry and Feldman（1985）會說：假如無法獲得更多的樣本資料，那只好「承認它的存在，接納它的後果」，尤其研究目的僅在作 R² 之預測時。



	x1	x2	x1x2	y
1	.00	1.00	.00	10.00
2	.50	1.00	.50	22.00
3	1.00	1.00	1.00	18.00
4	.00	2.00	.00	12.00
5	.50	2.00	1.00	6.00
6	1.00	2.00	2.00	8.00

圖 10 刪除共線性變項之資料檔

表 12 刪除共線性變項實例之相關矩陣

		y	x1	x2	x1x2
Pearson Correlation	y	1.000	0.145	-0.712	-0.257
	x1	0.145	1.000	0.000	0.885
	x2	-0.712	0.000	1.000	0.361
	x1x2	-0.257	0.885	0.361	1.000

由表 12 之相關矩陣知， $rx1x2=0$ ， $rx1x3=.885$ ， $rx2x3=.361$ 。

表 13 係全模式之迴歸分析結果。本模式中 x1 & x1x2 之 VIF 大於 10，出現多元共線性現象。

表 13 迴歸分析結果：全模式

Model	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
	B	Std. Error	Beta	T	Sig.	Tolerance	VIF
(Constant)	14.667	10.541		1.391	0.299		
x1	20.000	16.330	1.454	1.225	0.345	0.100	10.000
x2	-2.000	6.667	-0.178	-0.300	0.792	0.400	2.500
x1x2	-12.000	10.328	-1.479	-1.162	0.365	0.087	11.500

由表 14 知，刪除 x2 之後，交互作用項之斜率跟前述未刪除變項模式中之斜率估計值並未出現顯著不同，可能係因 $rx2x3=.361$ 並不高所致。

表 14 迴歸分析結果：刪除 x2 之次模式

Model	Unstandardized Coefficients		Standardized Coefficients		Collinearity Statistics		
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
(Constant)	11.667	2.782		4.193	0.025		
x1x2	-14.400	5.452	-1.774	-2.641	0.078	0.217	4.600
x1	23.600	9.244	1.715	2.553	0.084	0.217	4.600

由表 15 知，刪除 x1 之後，交互作用項之斜率逼近於 0，跟前述未刪除變項模式中之斜率估計值具有顯著不同，可能係因 $rx1x3=.885$ 很高所致。因此，當 x1 & x3 ($X3=X1X2$) 相關過低時，刪除共線性變項法將是不正確的作法。

表 15 迴歸分析結果：刪除 x1 之次模式

Model	Unstandardized Coefficients		Standardized Coefficients		Collinearity Statistics		
	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
(Constant)	24.667	7.201		3.426	0.042		
x1x2	2.55E-017	3.528	0.000	0.000	1.000	0.870	1.150
X2	-8.000	4.884	-0.712	-1.638	0.200	0.870	1.150

由此例可知，前述之多元共線性的解決方法大都具有應用上之限制，欲解決多元共線性的問題應針對資料的屬性對症下藥；解決方法當中，因為增大樣本蒐集更多更好資料，可以降低標準誤以減輕多元共線性的威脅（參見公式五之文字說明），似乎為最佳解決多元共線性策略。Berry & Feldman（1985, 1993）亦認為多元共線性乃是起因於樣本資訊不足所致，獲取更多樣本資訊乃是解決多元共線性的最佳策略。當蒐集更多資料不可行、中心化變項無效、合併變項不合理、刪除變項會造成模式界定錯誤時，那只好誠如 Berry & Feldman（1985）所言：「當取得更多的資訊不可行而又面臨嚴重的多元共線性時，最合理的途徑乃是：承認它的存在，接納它的後果（recognize its presence, but live with its consequence）」（p. 49）。

玖、結語

變項的中心化（或置中）的主要目的，在使截距的解釋具有實質意義與減輕多元共線性的威脅。變項間的高相關並不一定會出現多元共線性現象，而變項間的低相關也並不一定不會出現多元共線性現象，尚有其他因素（如資料結構屬性、估計誤差變異量）也會使得統計分析產生多元共線性現象。

本文以理論推導與實例說明了 VIF、Tolerance 指標旨在反映預測變項間之相關程度，而 CI、變異比率、特徵值指標則在反映預測變項的離散量。多元共線性的診斷不能單靠相關係數、VIF、Tolerance 指標，其它離散量的指標如 CI 與變異比率亦須同時全面診斷之。多元共線性的問題涉及許多重要因素：樣本大小、預測變項之變異量、變項間之相關性、極端值、整體變項之預測力與估計誤差變異量，均會造成參數估計標準誤過大而導致參數估計值不穩定現象。

變項的中心化只會影響交互作用項或高次方項的相關係數與 VIF 值；而且中心化並不一定會使交互作用項或高次方項的相關係數變小，有時也會變大（Shieh, 2011）。因此，中心化交互作用項有時亦會加重多元共線性的嚴重性，中心化並不能保證參數估計值不受多元共線性的威脅。此外，交乘積矩陣（ $X'X$ ）的行列式值在中心化與未中心化的資料上均完全相同，就行列式值來看，並無法改善多元共線性現象；有時也會因極端值而導致變項中心化，反而增強多元共線性的矛盾現象。中心化會減輕多元共線性之現象，主要原因係預測變項、交互作用項或高次方項（當作預測變項）的變異量常因相關變項之中心化而顯著下降所致。文中提出的「預測變項變異比差異」理論可以說明為何變項中心化，有時可以改善多元共線性的威脅，但有時卻不能改善的矛盾現象；尤其是交互作用項變異比乃是改善多元共線性的調節變項，當置中後其變異比如出現異常增大，中心化變項將無法有效降低原始變項與交互作用項間之多元共線性現象。

參數估計值不穩定的病因是多元的，因為多元共線性乃是樣本問題，其最佳方法是儘量蒐集更多更好的資料（Arceneaux & Huber, 2007）。因此，無任何單一統計方法可以完全減輕多元共線性的症狀。研究者必須對症下藥，如僅藉著變項的中心化減輕多元共線性，有時恐怕也會白忙一場或幫倒忙。

參考文獻

- Arceneaux, K. & Huber, G. A. (2007). What to do (and not do) with multicollinearity in state politics research. *State Politics and Policy Quarterly*, 7(1), 81-101.
- Belsley, D. A. (1984). Demeaning conditioning diagnostics through centering. *American Statistician*, 38(2), 73-77.
- Belsley, D. A. (1991). *Conditioning diagnostics: Collinearity and weakdata in regression*. New York: John Wiley & Sons.
- Belsley, D. (1991). A guide to using the collinearity diagnostics. *Management*, 4, 33-50.
- Berry, W. D. & Feldman, S. (1985). *Multiple regression in practice*. London: Sage Publications.
- Berry, W. D. & Feldman, S. (1993). *Multiple regression in practice*. In *Regression Analysis* (ed. Michael Lewis-Beck). London: Sage.
- Bottomley, P. & Holden, S. (2001). Do we really know how consumers evaluate brand extensions? Empirical generalizations based on secondary analysis of eight studies. *Journal of Marketing Research*, 38(4), 494-500.
- Chennamaneni, P., Echambadi, R., Hess, J. D., Syam, N. (2008). How do you properly diagnose harmful collinearity in moderated regressions? Retrieved June 1, 2011 from the World Wide Web
<http://som.utdallas.edu/academicAreas/marketing/documents/Hess.pdf>
- Cronbach, L. J. (1987). Statistical tests for moderator variables: Laws in analyses recently proposed. *Psychological Bulletin*, 102(3), 414-417.
- Echambadi, R., Arroniz, I., Reinartz, W., & Lee, J. (2006). Empirical generalizations from brand extension research: How sure are we? *International Journal of Research in Marketing*, 23(3), 253-261.
- Echambadi, R., & Hess, J. D. (2007). Mean-Centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Science*, 26(3), 438-445.
- Gatignon, H., & Vosgerau, J. (2005). *Moderating effects: The myth of mean centering*. Working paper. Retrieved June 1, 2011 from the World Wide Web
<http://www.insead.edu/facultyresearch/research/doc.cfm?did=1583>
- Hofmann, D. A. & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, 23, 723-744.
- Hofmann, D. A. (2007). Hierarchical linear modeling. Retrieved March 12, 2011 from the World Wide Web
http://www.unc.edu/~dhofmann/academy_2007_hlm_workshop.ppt#259,1, 投影片 1
- Judge, G. G., Griffiths, W. E., Hill, E. C., Lutkepohl, H. & Lee, T. C. (1985). *The Theory and Practice of Econometrics*, 2nd edn. New York: Wiley.
- Kreft, I. G., De Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, 30, 1-21.
- Lance, C. E. (1988). Residual centering, exploratory and confirmatory moderator analysis, and decomposition of effects in path models containing interactions. *Applied Psychological Measurement*, 12(6), 163-175.
- Lazaridis, A. (1986). A note regarding the problem of perfect multicollinearity.

- Quality and Quantity*, 120, 297–306.
- Liao, D. (2010). Collinearity diagnostics for complex survey data. Doctor of philosophy, joint program in survey methodology, University of Maryland. Retrieved June 1, 2011 from the World Wide Web
http://drum.lib.umd.edu/bitstream/1903/10881/1/Liao_umd_0117E_11537.pdf
- Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the merits of the orthogonalizing powered and product terms: Implications for modeling latent variable interactions. *Structural Equation Modeling*, 13, 479-519.
- Morris, J. H., Sherman, J. D., & Mansfield, E. R. (1986). Failures to detect moderating effects with ordinary least squares-moderated multiple regression: Some reasons and a remedy. *Psychological Bulletin*, 99(2), 282–288.
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690.
- Shieh, G. (2011). Clarifying the role of mean centering in multicollinearity of interaction effects. *British Journal of Mathematical and Statistical Psychology*, 64, 1-12.
- Wissmann, M., Toutenburg, H., & Shalabh. (2007). Role of categorical variables in multicollinearity in the linear regression model. Technical report number 008, Department of Statistics, University of Munich. Retrieved June 1, 2011 from the World Wide Web http://epub.ub.uni-muenchen.de/2081/1/report008_statistics.pdf.
- Wu, Y. W. B., & Wooldridge, P. J. (2005). The impact of centering first-level predictors on individual and contextual effects in multilevel data analysis. *Nursing Research*, 54(3), 212-216.
- Yu, M. (1989). Detection and remedies of multicollinearity in multiple regression analysis. *Journal of Education and Psychology*, 12, 385-407.