果樹產量組成因素分析與應用

Application of Yield Component Analysis in Fruit Crops

張龍生

國立台灣大學園藝學系
（本論文先送審於中國園藝）

摘　　要

本報告敘述標準化淨迴歸係數的估算，並比較迴歸係數、相關係數、以及標準化迴歸係數的產量。標準化淨迴歸係數可用來比較迴歸方程式各個獨立變量的迴歸係數應變量的重要程度，故可用來分析複合性狀－產量。本報告利用S.Wright的天竺鼠爲實例說明路徑係數分析圖，藉以闡明各個因素對天竺鼠出生後33天體重的影響。對複合性狀組成因素認知愈清楚，所得到的路路徑係數分析圖將更爲準確。產量組成要素的分析即利用此分析系統估測產量組成因素對產量的重要性，育種家可利用產量組成因素的分析供做選種與雜交的憑藉來達成產量育種的目的。

一、前言

單位面積產量的增加可藉由(1)改變植株的株型（Plant architecture），增加單位面積的栽植密度（Plant density)使每公頃的結果量達到最大。株型的改變可利用(A)矮生型（Dwarf)（B)密生型(Compact)（C）停心型（Determinate)等遺傳基因的調控直接降低植株高度與大小。所以有效結果面更密，更多的結實，達成密植栽培以增加產量。(2）假若作物果實內生性抑制劑造成座果率下降，可利用單爲結果（Parthenocarpic）的無籽化提高座果率；或全雌花性單爲結果的品種來提高產量（Baker等，1973)，Ghaderi和Lower(1981)等人曾育成全雌花單爲結果的胡瓜，達成提高單位面積的產量。(3）利用傳統族群的改良式，直接選拔影響產量組成的因素來提高產量。Blake(1943）曾經指出果樹產量的多寡受到遺傳組合

的控制，品種產量的潛能受到基因與外界環境交感作用的影響；產量是一個複雜的性狀，它不像前述兩者矮生型、停心型等性狀般的穩定，因矮生型與停心型可能受到某些主效因子的調控；產量是受到許多微效因子調控，需要有一正確的育種系統與方式來提高作物的產量。要成功的選拔良好基因型達到增產的目的，需要一套有效率的方式來估計產量，而且育種族群內須存在有廣泛的遺傳雜異度（genetic diversity）。當產量爲育種的目標，首先須對產量有所認知：(1) 產量是一個複雜的性狀。(2) 產量是數量遺傳性狀；各產量組成因素由微效基因所控制，易受環境影響。育種學家須有一合於選擇產量方式，在多動的環境下有效率的選擇所需要的良好基因型。然而各個產量組成因素彼此間對產量的影響多大呢？以往的研究常用相關係數來表示兩個變數間的關係，但並不能表示相對的重要性；本報告藉用路徑係數分析法來說明產量組成要素對產量的影響；作物育種家可利用這方析分式來判定各個產量組成因素對產量的相對影響，並可當作選種或雜交材料基因型選擇的憑藉。

二、相關係數與迴歸係數

科學試驗常將所有可能變異消除後，再研究一試驗條件對另一個狀況的影響。但造成變異的因素常非人爲所能操控，例如產量在生物研究上經常受一群因素的影響，而且各因素間還彼此存在交感作用；相關係數（coefficient of correlation）是最常用來表示兩個變因（Variable）的關係，相關係數利用前提乃是兩個變數間存有一直線關係；意指當某一個變因給予某程度的改變會造成另一變因某一程度的改變；它是一種定量的方式來描述兩個變因間的密切程度。相關係數的值介於-1與+1間；當 r=1時，表示完全正相關，當r=-1時表示完全負相關。

$$r=\frac{\dfrac{\Sigma(X-\overline{X})(Y-\overline{Y})}{n-1}}{\left[\dfrac{\Sigma(X-\overline{X})^2}{n-1}*\dfrac{\Sigma(Y-\overline{Y})^2}{n-1}\right]^{1/2}}$$

$$r=\frac{\Sigma(X-\overline{X})(Y-\overline{Y})}{\left[\Sigma(X-\overline{X})^2*\Sigma(Y-\overline{Y})^2\right]^{1/2}}$$

$$r=\frac{\Sigma XY}{(\Sigma X^2\ \Sigma Y^2)^{1/2}}$$

$$r=\frac{\Sigma XY}{\sigma_X\ \sigma_Y}$$

式中$\Sigma(X-\overline{X})(Y-\overline{Y})$稱為X與Y之離均差乘積和。

相關係數的平方即為判定係數（R^2），即總平方和中，由迴歸平方和所決定的部分，故當其中一個變數為常數時，$R^2=0$。故 R^2是用來估計變數Ｙ改變後所造成Ｘ改變程度，反之亦然。例如光合產物(Pn)與產量的相關係數r＝0.5；判定係數為0.25；表示兩者間彼此的變異程度只佔25％。故75％的其中一個的變異不受到另外一個因素的影響。相關係數的作用在於側定兩變數的相關程度，但並不能解釋兩變數間相互關係性質，例如主從關係，因果關係等。

$$R^2=\frac{(\Sigma XY)^2}{\Sigma X^2\ \Sigma Y^2}=\frac{\text{Reg}\quad S\cdot S}{\text{Tot}\quad S\cdot S}$$

迴歸係數（coefficient of regression b）常用來概括推論兩個變數在數量上的互變規律；兩個變量間存在最簡單的關係即是線性關係，我們常用Ｙ＝a＋bＸ表示；Ｙ通稱應變數（dependent variable），Ｘ通稱自變數（independent variable）。線性迴歸中， Y族群由相對應的 X值來決定，所以 Y的族群是一種逢分佈，也是一種常態分佈。迴歸係數即是斜率，意指 X每增加或減少一個單位時 Y平均地將要增加或減少的單位數；迴歸線代表兩個變數的關係與實際數值的誤差比率比其他任何直線都要小。故迴歸線是一條對各個分散點（data point）配合最好的直線，所以$\Sigma(Y-\overline{Y})=0$而且$\Sigma(Y-\overline{Y})^2$值最小。

$$byx=\frac{\Sigma xy}{\sigma_x}$$

迴歸係數是具有單位的數值，迴歸方程式常用來預測。但相關係數是沒有單位的數值只表示兩個變數間相關程度，不能作為預測用途。故產量組成分析各個產量因素間變異導致其它因子間的變異性質無法由相關係數得到，也不能用迴歸係數的大小來辨別各個變量對產量的重要程度。為了表示複雜性狀的各個變量的相關性質 S. wright於1921發表的路徑係數的分析，利用路徑途表示生長發育過程中各個變量的關係及其對複合性狀的重要程度。

三、產量組成因素分析

S. wright 於1921利用天竺鼠爲例考量兩個變量的因果關係。小天竺鼠出生33天斷奶後的重量＝小天竺鼠出生時產量＋出生33天所增加的體重。

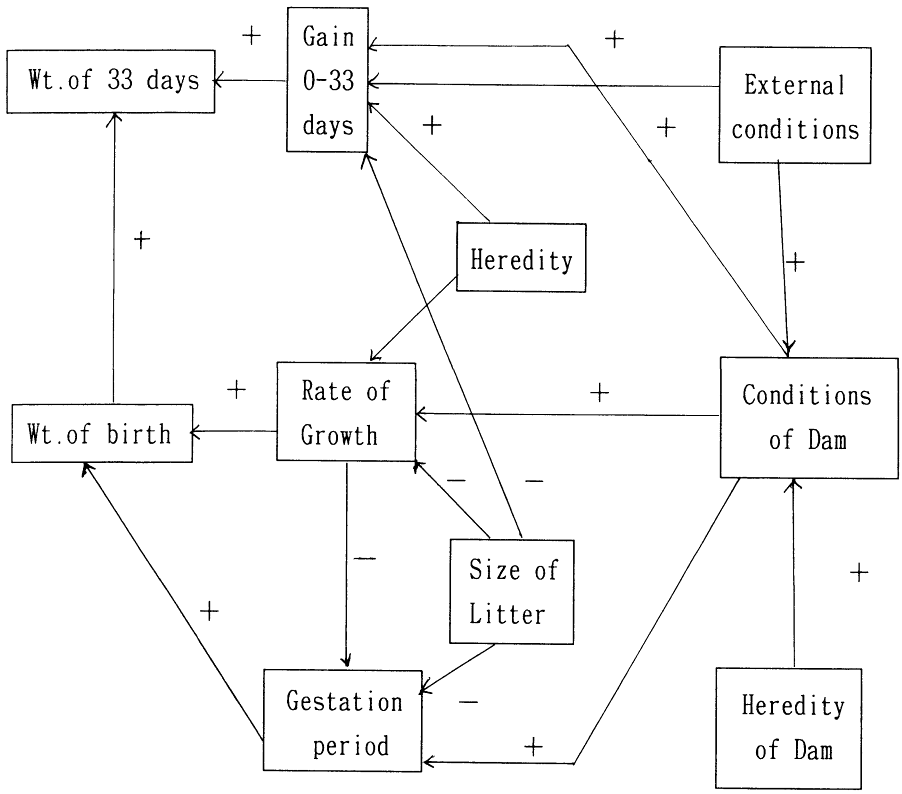

圖１　因素間交感作用對天竺鼠出生33天斷奶重量的路徑圖

天竺鼠出生的重量受到鼠體內妊娠時期的長短與母天竺鼠生長速率所決定，兩者對出生的重量呈現正相關關係；但母天竺鼠生長的快則會縮短母天竺鼠妊娠期，故爲負相關。妊娠長短又受母天竺鼠體內懷有的幼鼠的生長。鼠子愈多，生長速率慢，故爲負相關 。出生33天所增加的體重受鼠子的數目，遺傳質，母鼠狀況和外在環境條件的影響。母鼠的狀況先受到外界環境以及母鼠遺傳質的影響。所以路徑圖顯示了發育生長過程中前後優先順序。從圖可知，母鼠的生長速率會影響妊娠長短，但兩者間呈現數學關係而非因果關係；小天竺鼠出生時的體重可以由母天竺鼠懷孕期體重的增加的趨勢來決定。圖中也顯示鼠子出生時體重與33天斷奶後的體重基本上受遺傳質，同窩天竺鼠的鼠子數目，外在環境因子對母鼠懷孕期健康

狀況等因素的影響。S. Wright 首先利用路徑係數圖來表達層次間的關係，例如較大的生活的空間有利於母天竺鼠懷孕期的健康，會有較長的懷孕期以及較良好的生長狀況間接影響鼠子的出生時的重量。從圖中可瞭解一複合性狀表現受到各個因素的交互構成的網路影響，其表現極致就是結果的複合性狀。因而路徑係數即爲估量各個相關因素的直接效用的方法。今舉例來說明變量間的相關係數，迴歸係數與路徑係數。

表一　X與Z變量對Y呈直線關係

Y	X	Z
4.36	17	3
3.20	11	5
5.62	18	6
6.26	18	8
4.82	10	11
5.48	9	14
9.10	20	15
9.16	17	18
$\Sigma Y = 48$	$\Sigma X = 120$	$\Sigma Z = 80$
$\overline{Y} = 6$	$\overline{X} = 50$	$\overline{Z} = 10$

$\sigma_y^2 = 4$　　$\sigma_x^2 = 16$　　$\sigma_z^2 = 25$

$\sigma_y = 2$　　$\sigma_x = 4$　　$\sigma_z = 5$

迴歸方程式爲 $Y = -1.7 + 0.3X + 0.32Z$

$$\sigma_y^2 = \frac{\Sigma(Y-\overline{Y})^2}{n}$$

$$\sigma_{yx} = \left(\frac{\Sigma(Y-\overline{Y})(X-\overline{X})^{1/2}}{n}\right)$$

$$r_{yx} = \frac{\sigma_{yx}}{\sigma_y \ \sigma_x} = 0.6$$

$\sigma_{yz} = 8.0$

$$r_{yz} = \frac{\sigma_{yz}}{\sigma_y \ \sigma_z} = 0.8$$

$\sigma_{xz} = 0$

$r_{xz} = 0$

表二

Y	X	Z
2.26	10	3
2.60	9	5
6.22	20	6
5.96	17	8
5.12	11	11
8.18	18	14
8.50	18	15
9.16	17	18
$\Sigma Y=48$	$\Sigma X=120$	$\Sigma Z=80$
$\overline{Y}=6$	$\overline{X}=50$	$\overline{Z}=10$
$\sigma_y^2=5.92$	$\sigma_x^2=16$	$\sigma_z^2=25$
$\sigma_y=2.4331$	$\sigma_x=4$	$\sigma_z=5$

$\sigma_{yx}=8.0$

$$ryz=\frac{\sigma_{yx}}{\sigma_y\ \sigma_x}=0.8220$$

$\sigma_{yz}=11.0$

$r_{xz}=0.9042$

$\sigma xz=10$

迴歸方程式爲$Y=-1.7+0.3X+0.32Z$

兩個例子中得到迴歸方程式相同，故可以寫成

$$Y-\overline{Y}=B(X-\overline{X})+C(Z-\overline{Z})$$

$$\frac{(Y-\overline{Y})}{\sigma}=\frac{\sigma_X}{\sigma_y}-\frac{B(X-\overline{X})}{\sigma_X}+\frac{\sigma_Z}{\sigma_y}-\frac{C(Z-\overline{Z})}{\sigma_z}$$

$$=\frac{B\sigma X}{\sigma_y\sigma_X}(X-\overline{X})+\frac{C\sigma Z}{\sigma_y\ \sigma_Z}-(Z-\overline{Z})$$

$$若\ x=\frac{x-\overline{x}}{\sigma_x},\ y=\frac{y-\overline{y}}{\sigma_y},\ z=\frac{z-\overline{z}}{\sigma_z}$$

$$y=\frac{B\sigma_x}{\sigma_y}y+\frac{C\sigma_z}{\sigma_y}z$$

$y=Pyx\cdot X+Pyz\cdot Z$

$y=bx+cz$

表1的數據求得到標準化的淨迴歸係數 b與 c分別爲

$$b=\frac{B\sigma_x}{\sigma_y}=\frac{0.3*4}{2}=0.6$$

$$c=\frac{c\sigma_z}{\sigma_y}=\frac{0.32*5}{2}=0.8$$

$y=0.36x+0.8z$

表2的數據求得到標準化的淨迴歸係數 b與 c分別爲

$$b=\frac{B\sigma_x}{\sigma_y}=\frac{0.3*4}{2.4331}=0.4932$$

$$c=\frac{c\sigma_z}{\sigma_y}=\frac{0.32*5}{2.4331}=0.6576z$$

$y=0.4932x+0.6576z$

表1與表2的數據所得到標準化迴歸方程與直線迴歸方程不同。而且從公式上可知標準化的淨迴歸係數(standardized partial regression coefficient)沒有單位，故可用來比較兩個變量對複合性狀的重要性。標準化的淨迴歸係數即是路徑係數；而且可將相關係數劃分成直接效用與間接效用。

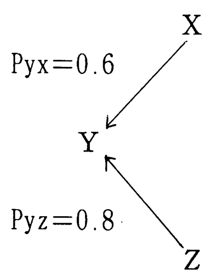

表1的數據

ryx＝Pyx＋rxzPyz＝0.6＋0*0.8＝0.6
ryz＝Pyz＋rxzPyx＝0.8＋0*0.6＝0.8

表2的數據

ryx＝Pyx＋rxzPyz＝0.4932＋0.5*0.6576＝0.8220
ryz＝Pyz＋rxzPyx＝0.65760＋0.5*0.4932＝0.9042

假若x與z兩個因素決定了y族群則

$$\sigma_y^2 = B^2\sigma_x^2 + c^2\sigma_z^2 + 2BC\sigma_x\sigma z\ rxz$$

兩邊同時除σ_y^2

$$\frac{\sigma_y^2}{\sigma_y^2} = \left(\frac{B\sigma_x}{\sigma_y}\right)^2 + \left(\frac{c\sigma_z}{\sigma_y}\right)^2 + 2\frac{B\sigma_x}{\sigma_y} + \frac{c\sigma_z}{\sigma_y} * rxz$$

$1 = P^2yx + Pyz + 2Pyx\ Pyz * rxz$
$1 = b^2 + c^2 + 2bcrxz$

例2的數據資料 $(0.4932)^2 + (0.6576)^2 + 2*0.4932*0.6576*0.5 = 1$

路徑係數分析方法應用於園藝界的實例中最明顯即是產量，(chgan et al,1986,1987)。因爲產量是一個複合性狀，產量是由產量組成因素組成，產量組成因素在作物生長發育期有層次順序，故可利用路徑係數分析產量與其組成因素的關係。故路徑係數的路徑分析圖是存有因果關係以及層次關係、而非任意建立的；欲建立一良好的路徑分析圖來闡明產量組成因素對產量的關係須植基於研究者對各產量因素層次順序以及生長生理的瞭解而認知組成因素的因果關係分析圖。若對變量間實質的關係認知愈清楚，所得到路徑圖來分析產量將更爲準確。

四、路徑係數分析於產量因素分析的應用

S. wright 於1921年發表路徑係數的分析（path analysis)，直到1956，他的學生C.C.

Li介紹，才廣爲育種學家所採納，應用於產量組成因素分析。藉由路徑係數分析，比較產量的組成要素間的差異以及對產量影響的程度。實質上，利用產量組成要素來選擇產量的方式，在1923年已由Engledow和Eadham提出，他們首先發表玉米產量：

產量＝單穗玉米粒數＊玉米粒平均重量

他們在報告建議玉米育種過程中選擇產量組成因素較直接選擇產量更能加速提高產量。然而他們並未發現Wright所發表路徑係數分析，利用路徑係數分析比較產量組成因素間的差異及其對產量的重要性。Williams利用產量組成因素的交感作用來解釋雜種優勢，他認爲雜種優勢的原因並主效因子（major gene）所造成的顯性或超顯性的結果。Williams以爲產量並非一個簡單的性狀，也非一個超級基因（super gene）所操控的產量；產量是由產量組成因素間的交感作用的結果，Williams是第一個提出雜種產量優勢非主效基因所控制而是產量組成因素間的交感作用，例如如F1的產量由果實數目與果實重量的交感作用而得到，故F_1呈現雜種優勢。

	Fruit number	Fruit weight	yield
P_1	1	3	3
P_2	3	1	3
F_1	2	2	4

Leng(1962)估算玉米產量與產量組成因素分析時有三個重要發現：(1）玉米的產量顯現了雜種優勢。(2）玉米粒的雜種數目優勢高於玉米粒重量的雜種優勢。(3）產量的遺傳率低，但產量組成因素的遺傳率高。

這些遺傳育種的數據中，若用路徑係數分析法來研判各個產量因素對產量影響程度，可提供育種家在育種上對產量組成性狀選擇的憑藉；再從育種族群內選出具某些優良產量組成因素的基因型供做父母本，再進行雜交；故產量組成因素間分析可提高產量育種選擇的效率。

各個產量組成因素間存在有正負相關性，產量組成因素間若呈現負相關的情況，W. Adams稱爲產量補償作用（yield compensation）。Adams推論補償作用乃起因於兩個產量組成因素間彼此競爭代謝產物，產量組成因素系統受到限制而導致發育受阻而弔起負相關。例如天竺鼠例子中母天竺鼠妊娠時期與母天竺鼠的生長速率呈現負相關。母鼠生長愈快，妊娠時間短，則造成鼠子出生時的重量減少，但母天竺鼠速率又受外界環境所影響。產量組成因素受外界環境影響，外界環境可引起植物植物個體彼此的競爭會導致植物體內的競爭，而構成

產量補償作用。

五、結論

(1)標準化的淨迴歸係數可用來比較各個迴歸係數，而且可用來表示各個變量的相關性質。

(2)路徑係數分析用來估量產量組成因素彼此間因果關係的相關重要性。

(3)產量組成因素選拔較直接對產量選擇更為有效率。

(4)育種家可利用產量組成因素分析供做育種族群良好基因型的選擇依據。

(5)產量組成因素間遺傳背景瞭解，以及各個因素間層次次序瞭解利有利用建立完善的產量組成因素分析圖。

Summary

(1) Standarrdized partial regression coefficient is the direct path coefficient, which can be used to determine the relative.

(2) Path coefficient analysis can be used to evaluate the relationships between yield components using causatine relationships.

(3) Selection for individual yield components is more efficient than selecting for yield itself.

(4) Plant breeders can seperate yield into the product of its part and choose parents selected for their independent yield components superiorities.

(5) More information is needed concerning the genetic control of yield components, especially for fruit tree crop breeding.

六、參 考 文 獻

(1) Adams,M.W. 1967. Basis of yield component compensation in crop plants with special reference to the field bean, Phaseolus vulgaris,Crop Sci.7:505-510.

(2) Adams,M.W. and J.E. Grafius 1971. Yield component compensationalternative interpretation, Crop Sci. 11:33-15.

(3) Blake, M.A. 1943 Classification of fruit bud development on peaches and nectarines and its significance in cultural practice, N.J. Agr. Expt. Sta. bull. 706.

(4) Baker, L.R, J.W. Scott, and J.E. Wilson, 1973. Seedless pickles--a new co- ncetp, Mich State Univ. Res. Rep 277.

(5) Chang, L.S., A.F. Iezzoni, and J.A. Flore. 1986. Variation in yield compo- nents between "Mont- morency" and "Meteor" sour eherry. Hort. Sci. 21:711.

(6) Chang, L.s., A.F. Iezzoni, and J.A. Flore. 1987. Yield component in "Mont- morency" and "Meteor" sour chrry. J. Amer. Soc. Hort. Sci. 112:247-251.

(7) El-Shawarf, I.I.S. and L.R. Barker. 1981. Performance of hermaphroditic in gynoecious pickling cu- cumber for one-ever mechanical harvest. J.Amer. soc. Hort. Sci. 106:356-359.

(8) Gjaderi, A. and R.L. Lower 1981. Estimates of genetic variance for yield in pickling cucumbers. J. Amer. Soc. Hort. Sci. 106:237-239.

(9) Leng, E.R. 1962. Component analysis in ingeritance studies of grain yield in maize. Crop Sci. 3:186-190.

(10) Li, C.C. 1956. The concept of path coefficients and its impact on Natare. 184:527-530.

(11) William, M.W. 1959. Heterosis and the genetics of complex characters. Natare. 184:527-530. (12) Wright, s 1921. Correlation and causation. J. Agron. Res. 20:557-587.