الفصل الثاني: التحليل ألعاملي العام

Analyse Factorielle Générale (AFG)

تمهيد

تعتبر طريقة التحليل ألعاملي العام أساس كل طرق تحليل المعطيات للمتغيرات الكمية، و الفكرة الأساسية للتحليل ألعاملي العام هي البحث عن فضاء شعاعي جزئي اقل درجة عادةً ما يكون ذو البعد 2 يسمح لنا بأحسن تمثيل للبيانات و يحفظ لنا اكبر كمية من المعلومات، و يكون هذا بالنسبة للأفراد و المتغيرات.

1. التحليل في $\Re^p$

ليكن $X(n,p)$ الجدول الأولي للمعطيات الذي يحتوي على n فرد ممثلين في الأسطر و p متغير أو خاصية ممثلة في الأعمدة، و نكتب:

$$X(n,p)=\begin{array}{c|c} & \begin{array}{cccccc} 1 & 2 & \ldots\ldots & j & \ldots\ldots & p \end{array} \\ \begin{array}{c} 1 \\ . \\ . \\ i \\ . \\ n \end{array} & \begin{pmatrix} x_{11} & x_{12} & .. & x_{1j} & .. & x_{1p} \\ .. & .. & .. & .. & \cdots & \\ .. & .. & .. & .. & .. & \\ x_{i1} & x_{i2} & .. & x_{ij} & .. & x_{ip} \\ .. & .. & .. & .. & & .. \\ x_{n1} & x_{n2} & .. & x_{nj} & .. & x_{np} \end{pmatrix} \end{array} \qquad i=\overline{1,n} \quad ; \quad j=\overline{1,p}$$

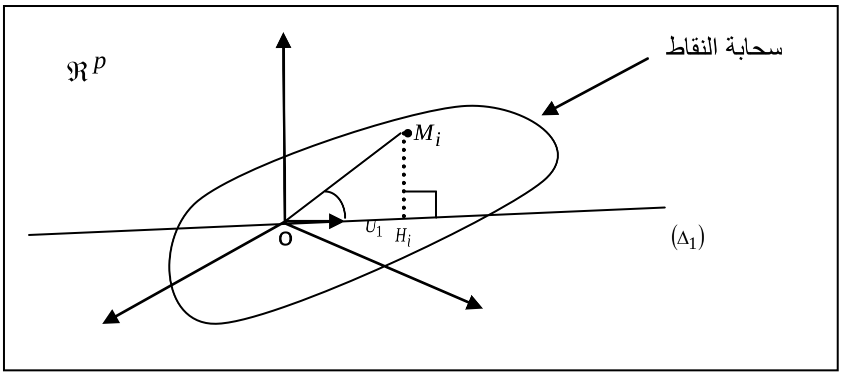
حيث أن: x_{ij} هو العنصر المولد للمصفوفة $X(n,p)$ و يمثل قيمة الفرد i عند المتغير j ؛

x_i هو شعاع سطر و يمثل كل قيم الفرد i ؛

x^j هو شعاع عمود و يمثل كل قيم المتغير j ؛

ان المقصود بالتحليل في $\Re^p$ هو دراسة و تحليل n فرد في فضاء المتغيرات ذو P محور، فالتمثيل البياني لـ n فرد في الفضاء $\Re^p$ يعطي لنا سحابة نقاط موضحة في الشكل التالي:

الشكل (1) : التمثيل البياني للتحليل في $\Re^p$



إن التمثيل البياني السابق هو صورة توضيحية لـ n فرد في الفضاء $\Re^p$ أي الفضاء الذي يحتوي على p محور، غير أن تحليل و دراسة التمثيل السابق لا يكون إلا على سطح الورقة أو سطح المكتب أي على فضاء شعاعي جزئي ذو البعد 2، و عليه فان الصورة السابقة تصبح غير واضحة المعالم و لا يمكننا ملاحظة أشياء كثيرة منها. و على أساس هذا التحليل فإننا نبحث عن فضاء شعاعي جزئي اقل درجة عادةً ما يكون ذو البعد 2 (سطح الورقة أو سطح المكتب) يسمح لنا بأحسن تمثيل للبيانات.

غير أننا نعلم أن الانتقال من فضاء ذو بعد أعلى إلى فضاء جزئي اقل درجة يؤدي بنا إلى احتمال فقدان جزء من المسافات أي إمكانية فقدان جزء من المعلومات، وبالتالي فان البحث عن فضاء شعاعي جزئي اقل درجة يسمح لنا بأحسن تمثيل للبيانات و يحفظ لنا اكبر كمية من المعلومات.

فإذا اعتبرنا أن (Δ_1) هو احد محاور الفضاء الشعاعي الجزئي المراد البحث عنه، حيث أن (Δ_1) يخترق المبدأ O ، فان المسافة الحقيقية هي $[OM_i]$ أما $[OH_i]$ فتمثل مسافة إسقاط $[OM_i]$ على المحور (Δ_1) أي صورة المعلومات للفرد i على هذا المحور، و كما نعلم مسبقاً فان مسافة الإسقاط دوماً اقل أو تساوي المسافة الحقيقية و نكتب:

$$[OH_i] \leq [OM_i]$$

أي أن الانتقال من الفضاء الأصلي إلى المحور (Δ_1) يؤدي بنا إلى إمكانية فقدان جزء من المسافات أي إمكانية فقدان جزء من المعلومات، و أنت كباحث لا تحبذ فقدان المعلومات أي أننا نبحث عن المحور (Δ_1) الذي يقلل من فقدان المعلومات أي الذي يُعظم المسافة OH_i ، و على أساس أن OH_i هي إحداثية النقطة M_i على (Δ_1) فانه قد تكون القيمة OH_i سالبة وبالتالي يجب تعظيم القيمة OH_i^2 ، كما أننا نبحث عن المحور (Δ_1) الذي يقلل من فقدان المعلومات ليس فقط بالنسبة للنقطة i وإنما بالنسبة لكل النقط وبالتالي يجب تعظيم القيمة:

$$Q = \sum_{i=1}^{n} OH_i^2$$

ليكن U_1 شعاع توجيه المحور (Δ_1) أي أن:

$$U_1' U_1 = 1 \Rightarrow \begin{pmatrix} u_1 & u_2 & \dots & u_j & \dots & u_p \end{pmatrix} \otimes \begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_j \\ \cdot \\ \cdot \\ u_p \end{pmatrix} = \sum_{j=1}^{p} u_j^2 = 1$$

و إذا كان إسقاط النقطة M_i على المحور (Δ_1) يعطينا النقطة H_i ، فان البعد OH_i يمثل إحداثية النقطة i على المحور (Δ_1) ، و نكتب:

$$OH_i = x_i \times U_1$$

حيث أن x_i تمثل إحداثية الفرد i في الفضاء $\Re^p$:

$$x_i = \left(x_{i1};\quad x_{i2};\quad ....\quad x_{ij};\quad ....\quad ; x_{ip}\right)$$

نرمز بـ F_1 لإحداثيات كل الأفراد على المحور (Δ_1) و نكتب:

$$F_1 = XU_1 = \begin{pmatrix} OH_1 \\ OH_2 \\ . \\ . \\ OH_i \\ . \\ . \\ OH_n \end{pmatrix} = \begin{pmatrix} F_1(1) \\ F_1(2) \\ . \\ . \\ F_1(i) \\ . \\ . \\ F_1(n) \end{pmatrix}$$

فالقيمة $OH_1 = F_1(1)$ تمثل إحداثية الفرد الأول على المحور (Δ_1) و القيمة $OH_2 = F_1(2)$ تمثل إحداثية الفرد الثاني و القيمة $OH_n = F_1(n)$ تمثل إحداثية الفرد n على نفس المحور (Δ_1) .

و بغرض كتابة القيمة Q على شكل مصفوفاتي فانه لدينا:

$$F_1{}' F_1 = \left(OH_1 \quad OH_2 \quad .... \quad OH_i \quad .... \quad OH_n\right) \times \begin{pmatrix} OH_1 \\ OH_2 \\ . \\ . \\ OH_i \\ . \\ . \\ OH_n \end{pmatrix} = \sum_{i=1}^{n} OH_i^2 = Q$$

وعليه يمكننا أن نكتب:

$$Q = \sum_{i=1}^{n} OH_i^2 = F_1{}' F_1 = (XU_1)'(XU_1) = U_1{}' X' X U_1$$

ويصبح المشكل المطلوب حله هو:

$$\begin{cases} MAX \;\; Q = \sum_{i=1}^{n} OH_i^2 = U_1{}' X' X U_1 \\ S/C \\ U_1{}' U_1 = 1 \end{cases}$$

و بغرض حل المشكل أعلاه فإننا نستعمل دالة لاغرانج:

$$L(U_1, \lambda_1) = U_1{}' X' X U_1 - \lambda_1 \left(U_1{}' U_1 - 1\right)$$

وحتى نحصل على القيمة العظمى للدالة Q يجب أن يكون مشتق دالة لاغرانج $L(U_1,\lambda_1)$ بالنسبة لـ U_1 و بالنسبة لـ λ_1 معدوم و نكتب:

$$\begin{cases} \dfrac{\partial L(U_1,\lambda_1)}{\partial U_1}=0 \\ \dfrac{\partial L(U_1,\lambda_1)}{\partial \lambda_1}=0 \end{cases} \Rightarrow \begin{cases} 2X'XU_1-2\lambda_1 U_1=0 \\ U_1{}'U_1-1=0 \end{cases} \Rightarrow \begin{cases} X'XU_1=\lambda_1 U_1 \ldots\ldots\ldots\ldots(1) \\ U_1{}'U_1=1 \ldots\ldots\ldots\ldots\ldots(2) \end{cases}$$

من المعادلة (1) يمكننا القول أن U_1 هو شعاع ذاتي للمصفوفة $(X'X)$ و المرفق بالقيمة الذاتية λ_1 .

غير أننا نعلم أن المصفوفة $\left(X'_{p\times n}X_{n\times p}\right)_{p\times p}$ هي مصفوفة مربعة ذات p سطر و p عمود تملك p قيمة ذاتية، فأي قيمة ذاتية نأخذ؟ من المعادلة (1) يمكننا أن نكتب:

$$(1)\Rightarrow X'XU_1=\lambda_1 U_1 \Rightarrow U_1{}'X'XU_1=U_1{}'\lambda_1 U_1=\lambda_1 U_1{}'U_1$$

و من المعادلة (2) لدينا: $U_1{}'U_1=1$ و بالتعويض في العبارة أعلاه نجد أن:

$$Q=U_1{}'X'XU_1=\lambda_1$$

و على أساس أن المشكل هو تعظيم القيمة $Q=U_1{}'X'XU_1$ و وجدنا أن $Q=\lambda_1$ فإننا نأخذ اكبر قيمة ذاتية λ_1 من بين كل القيم الذاتية المستخرجة للمصفوفة $(X'X)$ كقيمة تسمح لنا باستخراج الشعاع الذاتي U_1 و الذي يعتبر شعاع توجيه للمحور (Δ_1) و الذي يعتبر أحسن محور على الإطلاق يحفظ لنا اكبر كمية من المعلومات و يعطينا أحسن تمثيل للبيانات، وعليه فإننا نقوم بتحليل المصفوفة $(X'X)$ و استخراج p قيمة ذاتية ثم نرتبها ترتيباً تنازلياً القيمة الأكبر ثم التي تليها و بعد ذلك نجد الأشعة الذاتية المرافقة لها و نكتب:

القيم الذاتية: $\lambda_1;\lambda_2;\lambda_3\ldots\ldots\ldots\ldots\ldots;\lambda_p$

الأشعة الذاتية: $U_1;U_2;U_3\ldots\ldots\ldots\ldots\ldots;U_p$

المحاور: $\Delta_1;\Delta_2;\Delta_3\ldots\ldots\ldots\ldots\ldots;\Delta_p$

فالمحور (Δ_1) ذو شعاع توجيه U_1 يعتبر أحسن محور على الإطلاق يحفظ لنا اكبر كمية من المعلومات و يعطينا أحسن تمثيل للبيانات، ثم المحور (Δ_2) ذو شعاع توجيه U_2 ثم المحور الذي يليه و هكذا بهذا الترتيب. وعليه فان المستوي الأول $(\Delta_1\otimes\Delta_2)$ المشكل من المحورين الأول (Δ_1) و الثاني (Δ_2) هو المستوي الأفضل على الإطلاق يعطينا أحسن تمثيل و يحفظ لنا اكبر كمية من المعلومات.

1.1. الكثافة الكلية للبيانات

إن الكثافة الكلية للبيانات هي مجموع مربعات بعد كل النقاط عن المبدأ (O) أي بمعنى آخر هي التشتت الكلي بالنسبة للمبدأ، و نكتب:

$$It = \sum_{i=1}^{n} d(o,i)^2 = \sum_{\alpha=1}^{p} F_\alpha{}' F_\alpha = \sum_{\alpha=1}^{p} \lambda_\alpha = Tr(X'X)$$

أي أن مجموع القيم الذاتية يمثل الكثافة الكلية للبيانات و التي تعبر عن نوعية التمثيل أو التقديم البياني، و يمكننا كذلك أن نسمي الكثافة الكلية للبيانات بالجمود الكلي.

2.1. نسب تمثيل البيانات على المحاور

إن الانتقال من الفضاء الأصلي إلى المحور (Δ_α) يؤدي بنا إلى فقدان جزء من المعلومات و اكتساب جزء من المعلومات على هذا المحور، فالنسبة المكتسبة على هذا المحور هي نسبة تمثيل البيانات على المحور، و تعطي هذه النسبة بالعلاقة التالية:

$$\tau_\alpha = \frac{\lambda_\alpha}{It} = \frac{\lambda_\alpha}{\sum_{\alpha=1}^{p} \lambda_\alpha}$$

أي أن $\tau_\alpha \%$ من بيانات المصفوفة X ممثلة على المحور (Δ_α).

فنسبة التمثيل على المحور الأول هي: $\tau_1 = \frac{\lambda_1}{It}$

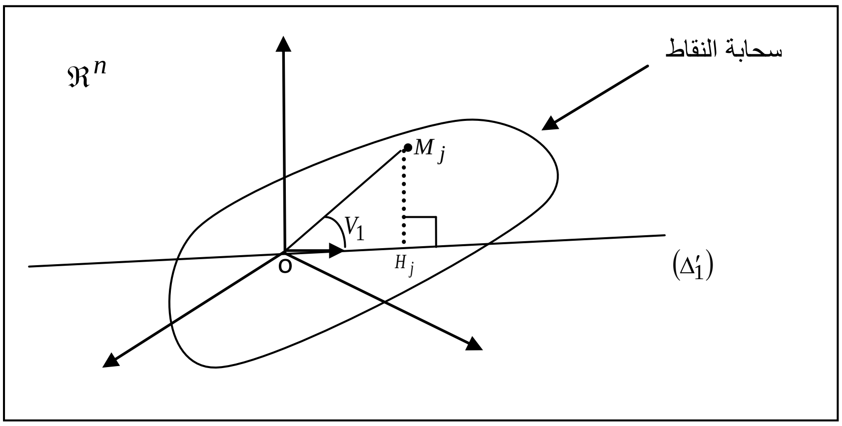
فنسبة التمثيل على المحور الثاني هي: $\tau_2 = \frac{\lambda_2}{It}$

فتكون نسبة التمثيل على المستوي الأول $(\Delta_1 \otimes \Delta_2)$ هي: $\tau(\Delta_1 \otimes \Delta_2) = \frac{\lambda_1 + \lambda_2}{It}$

2. التحليل في $\Re^n$

كما هو معلوم سابقاً فان المقصود بالتحليل في $\Re^n$ هو دراسة و تحليل p متغير في فضاء الأفراد ذو n محور، فالتمثيل البياني لـ p متغير في الفضاء $\Re^n$ يعطي لنا سحابة نقاط موضحة في الشكل التالي:

الشكل (2) : التمثيل البياني للتحليل في $\Re^n$



إن التمثيل البياني السابق هو صورة توضيحية لـ p متغير في الفضاء $\Re^n$ أي الفضاء الذي يحتوي على n محور، غير أن تحليل و دراسة التمثيل السابق لا يكون إلا على سطح الورقة أو سطح المكتب أي على فضاء شعاعي جزئي ذو البعد 2، و عليه فان الصورة السابقة تصبح غير واضحة المعالم و لا يمكننا ملاحظة أشياء كثيرة منها. و على أساس هذا التحليل فإننا نبحث عن فضاء شعاعي جزئي اقل درجة عادةً ما يكون ذو البعد 2 (سطح الورقة أو سطح المكتب) يسمح لنا بأحسن تمثيل للمتغيرات، غير أننا كما نعلم مسبقاً أن الانتقال من فضاء ذو بعد أعلى إلى فضاء جزئي اقل درجة يؤدي بنا إلى احتمال فقدان جزء من المسافات أي إمكانية فقدان جزء من المعلومات، وبالتالي فان البحث عن فضاء شعاعي جزئي اقل درجة يسمح لنا بأحسن تمثيل للبيانات و يحفظ لنا اكبر كمية من المعلومات.

فإذا اعتبرنا أن (Δ_1') هو احد محاور الفضاء الشعاعي الجزئي المراد البحث عنه، حيث أن (Δ_1') يخترق المبدأ 0 ، فان المسافة الحقيقية هي $|OM_j|$ أما $|OH_j|$ فتمثل مسافة إسقاط $|OM_j|$ على المحور (Δ_1') أي صورة المعلومات للمتغير j على هذا المحور، و كما نعلم مسبقاً فان مسافة الإسقاط دوماً اقل أو تساوي المسافة الحقيقية و نكتب:

$$|OH_j| \leq |OM_j|$$

أي أن الانتقال من الفضاء الأصلي إلى المحور (Δ_1') يؤدي بنا إلى إمكانية فقدان جزء من المسافات أي إمكانية فقدان جزء من المعلومات، و أنت كباحث لا تحبذ فقدان المعلومات أي أننا نبحث عن المحور (Δ_1') الذي يقلل من فقدان المعلومات وبالتالي الذي يُعظم المسافة OH_j ، و على أساس أن OH_j هي إحداثية النقطة M_j على (Δ_1') فانه قد تكون القيمة OH_j سالبة وبالتالي يجب تعظيم القيمة OH_j^2.

كما أننا نبحث عن المحور (Δ_1') الذي يقلل من فقدان المعلومات ليس فقط بالنسبة للنقطة j فقط وإنما بالنسبة لكل النقط وبالتالي يجب تعظيم القيمة: $Q' = \sum_{j=1}^{p} OH_j^2$

ليكن V_1 شعاع توجيه المحور (Δ_1') أي أن:

$$V_1' V_1 = 1 \Rightarrow \begin{pmatrix} v_1 & v_2 & \ldots & v_i & \ldots & v_n \end{pmatrix} \otimes \begin{pmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_i \\ \cdot \\ \cdot \\ v_n \end{pmatrix} = \sum_{i=1}^{n} v_i^2 = 1$$

و إذا كان إسقاط النقطة M_j على المحور (Δ_1') يعطينا النقطة H_j ، فان البعد OH_j يمثل إحداثية النقطة j على المحور (Δ_1') و نكتب: $OH_j = \left(x^j\right)' \times V_1$

حيث أن $\left(x^j\right)'$ تمثل إحداثية المتغير j في الفضاء $\Re^n$:

$$\left(x^j\right)' = \left(x_{1j}; \quad x_{2j}; \quad .... \quad x_{ij}; \quad .... \quad ; x_{nj}\right)$$

نرمز بـ G_1 لإحداثيات كل المتغيرات على المحور (Δ_1') و نكتب:

$$G_1 = X'V_1 = \begin{pmatrix} OH_1 \\ OH_2 \\ . \\ . \\ OH_j \\ . \\ . \\ OH_p \end{pmatrix} = \begin{pmatrix} G_1(1) \\ G_1(2) \\ . \\ . \\ G_1(j) \\ . \\ . \\ G_1(p) \end{pmatrix}$$

فالقيمة $OH_1 = G_1(1)$ تمثل إحداثية المتغير الأول على المحور (Δ_1') و القيمة $OH_2 = G_1(2)$ تمثل إحداثية المتغير الثاني و القيمة $OH_P = G_1(P)$ تمثل إحداثية المتغير p على نفس المحور (Δ_1') .

و بغرض كتابة القيمة Q' على شكل مصفوفاتي فانه لدينا:

$$G_1'G_1 = \left(OH_1 \quad OH_2 \quad .... \quad OH_j \quad .... \quad OH_p\right) \times \begin{pmatrix} OH_1 \\ OH_2 \\ . \\ . \\ OH_j \\ . \\ . \\ OH_p \end{pmatrix} = \sum_{j=1}^{p} OH_j^2 = Q'$$

وعليه يمكننا أن نكتب:

$$Q' = \sum_{j=1}^{p} OH_j^2 = G_1'G_1 = \left(X'V_1\right)'\left(X'V_1\right) = V_1'XX'V_1$$

ويصبح المشكل المطلوب حله هو:

$$\begin{cases} MAX \;\; Q' = \sum_{j=1}^{p} OH_j^2 = V_1'XX'V_1 \\ S/C \\ V_1'V_1 = 1 \end{cases}$$

و بغرض حل المشكل أعلاه فإننا نستعمل دالة لاغرانج:

$$L(V_1,\mu_1) = V_1{}' XX'V_1 - \mu_1\left(V_1{}'V_1 - 1\right)$$

وحتى نحصل على القيمة العظمى للدالة Q' يجب أن يكون مشتق دالة لاغرانج $L(V_1,\mu_1)$ بالنسبة لـ V_1 و بالنسبة لـ μ_1 معدوم و نكتب:

$$\begin{cases}\dfrac{\partial L(V_1,\mu_1)}{\partial V_1} = 0\\[2mm] \dfrac{\partial L(V_1,\mu_1)}{\partial \mu_1} = 0\end{cases} \Rightarrow \begin{cases}2XX'V_1 - 2\mu_1 V_1 = 0\\ V_1{}'V_1 - 1 = 0\end{cases} \Rightarrow \begin{cases}XX'V_1 = \mu_1 V_1 \ldots\ldots\ldots\ldots\ldots(1)\\ V_1{}'V_1 = 1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)\end{cases}$$

من المعادلة (1) يمكننا القول أن V_1 هو شعاع ذاتي للمصفوفة (XX') و المرفق بالقيمة الذاتية μ_1 . غير أننا نعلم أن المصفوفة $\left(X_{n\times p}X'_{p\times n}\right)_{n\times n}$ هي مصفوفة مربعة ذات n سطر و n عمود تملك n قيمة ذاتية، فأي قيمة ذاتية نأخذ؟ من المعادلة (1) يمكننا أن نكتب:

$$(1) \Rightarrow XX'V_1 = \mu_1 V_1 \Rightarrow V_1{}' XX'V_1 = V_1{}'\mu_1 V_1 = \mu_1 V_1{}'V_1$$

و من المعادلة (2) لدينا: $V_1{}'V_1 = 1$ و بالتعويض في العبارة أعلاه نجد أن:

$$Q' = V_1{}' XX'V_1 = \mu_1$$

و على أساس أن المشكل هو تعظيم القيمة $Q' = V_1{}' XX'V_1$ و وجدنا أن $Q' = \mu_1$ فإننا نأخذ اكبر قيمة ذاتية μ_1 من بين كل القيم الذاتية المستخرجة للمصفوفة (XX') كقيمة تسمح لنا باستخراج الشعاع الذاتي V_1 و الذي يعتبر شعاع توجيه للمحور $\left(\Delta_1{}'\right)$ و الذي يعتبر أحسن محور على الإطلاق يحفظ لنا اكبر كمية من المعلومات و يعطينا أحسن تمثيل للبيانات، وعليه فإننا نقوم بتحليل المصفوفة (XX') و استخراج n قيمة ذاتية ثم نرتبها ترتيباً تنازلياً القيمة الأكبر ثم التي تليها و بعد ذلك نجد الأشعة الذاتية المرافقة لها و نكتب:

القيم الذاتية: $\mu_1;\mu_2;\mu_3\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots;.\mu_n$

الأشعة الذاتية: $V_1;V_2;V_3\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots;.V_n$

المحور: $\Delta_1{}';\Delta_2{}';\Delta_3{}'\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots;.\Delta_n{}'$

فالمحور $\left(\Delta_1{}'\right)$ ذو شعاع توجيه V_1 يعتبر أحسن محور على الإطلاق يحفظ لنا اكبر كمية من المعلومات و يعطينا أحسن تمثيل للبيانات، ثم المحور $\left(\Delta_2{}'\right)$ ذو شعاع توجيه V_2 ثم المحور الذي يليه و

هكذا بهذا الترتيب. وعليه فان المستوي الأول $\left(\Delta_1' \otimes \Delta_2'\right)$ المشكل من المحورين الأول $\left(\Delta_1'\right)$ و الثاني $\left(\Delta_2'\right)$ هو المستوي الأفضل على الإطلاق يعطينا أحسن تمثيل و يحفظ لنا اكبر كمية من المعلومات.

3. العلاقة بين التحليل في $\Re^p$ و التحليل في $\Re^n$

من خلال نتائج التحليل في الفرعين السابقين وجدنا أن:

التحليل في $\Re^p$: $\alpha = \overline{1, p}$............(1) $\quad (X'X)\ U_\alpha = \lambda_\alpha U_\alpha$

التحليل في $\Re^n$: $\beta = \overline{1, n}$...........(2) $\quad (XX')\ V_\beta = \mu_\beta V_\beta$

في البداية نضرب المعادلة (1) من اليسار في المصفوفة X فنجد أن:

$$(1) \Rightarrow (X'X)U_\alpha = \lambda_\alpha U_\alpha \Rightarrow X(X'X)U_\alpha = X\lambda_\alpha U_\alpha \Rightarrow (XX')(XU_\alpha) = \lambda_\alpha (XU_\alpha)$$

و يمكننا أن نقرأ العبارة أعلاه أن الشعاع (XU_α) هو شعاع ذاتي للمصفوفة (XX') و المرفق بالقيمة الذاتية λ_α ، غير أننا نعلم من التحليل في $\Re^n$ أن μ_β هي القيمة الذاتية الأكبر للمصفوفة (XX') وعليه يكون: $\mu_\beta \geq \lambda_\alpha$

و من جهة أخرى يمكننا أن نضرب المعادلة (2) من اليسار في المصفوفة X' فنجد أن:

$$(2) \Rightarrow (XX')V_\beta = \mu_\beta V_\beta \Rightarrow X'(XX')V_\beta = X'\mu_\beta V_\beta \Rightarrow (X'X)(X'V_\beta) = \mu_\beta (X'V_\beta)$$

و يمكننا أن نقرأ العبارة أعلاه أن الشعاع $(X'V_\beta)$ هو شعاع ذاتي للمصفوفة $(X'X)$ و المرفق بالقيمة الذاتية μ_β ، غير أننا نعلم من التحليل في $\Re^p$ أن λ_α هي القيمة الذاتية الأكبر للمصفوفة $(X'X)$ وعليه يكون: $\lambda_\alpha \geq \mu_\beta$

و على أساس هاتين النتيجتين فان: $\mu_\beta \geq \lambda_\alpha$ و $\lambda_\alpha \geq \mu_\beta$ يكون من الضروري: $\lambda_\alpha = \mu_\beta$

كما أننا نعلم أن عدد الأفراد عادةً اكبر أو يساوي عدد المتغيرات: $n \geq p$ أي أن سلسلة قيم α اقل من سلسلة قيم β ، و بغرض تحقيق التطابق بين قيم λ_α و قيم μ_β يجب أن يكون طول قيم السلسلة الأولى مساوي لطول قيم السلسلة الثانية و مساوية للعدد الأقل و هو α و يحدث ذلك بفعل التطابق، و نكتب: $\alpha = \overline{1, p} \quad \lambda_\alpha = \mu_\alpha$

و من خلال النتيجة السابقة يمكننا أن نكتب:

التحليل في $\Re^p$: $\alpha = \overline{1, p}$............(1) $\quad (X'X)\ U_\alpha = \lambda_\alpha U_\alpha$

التحليل في $\Re^n$: $\alpha = \overline{1, p}$...........(2) $\quad (XX')\ V_\alpha = \lambda_\alpha V_\alpha$

4. عبارات الانتقال بين $\Re^p$ و $\Re^n$

و على أساس نتائج التحليل السابقة يمكننا أن نستنتج عبارات الانتقال التالية، وهذه العبارات تسمح لنا بالانتقال من تحليلٍ إلى آخر، و فائدة عبارات الانتقال هي تبسيط و تسهيل الحسابات و توفير الوقت.

1.4. عبارة أول انتقال من $\Re^p$ نحو $\Re^n$

إذا قمنا بالتحليل في $\Re^p$ أي أننا نملك القيم الذاتية λ_α و الأشعة الذاتية U_α و نرغب الآن في الانتقال إلى $\Re^n$ بغيت دراسة و تحليل المتغيرات، أي يجب علينا تحليل المصفوفة (XX') و استخراج القيم الذاتية λ_α و الأشعة الذاتية V_α.غير أننا من التحليل السابق وجدنا أن الشعاع (XU_α) هو شعاع ذاتي للمصفوفة (XX') و المرفق بالقيمة الذاتية λ_α ، يمكننا أن نضع مرحلياً: $V_\alpha = XU_\alpha$

و حتى يكون الشعاع الذاتي أعلاه شعاع توجيه للمحور α يجب أن تكون طويلة هذا الشعاع تساوي الواحد، غير أن طويلة هذا الشعاع هي:

$$\|V_\alpha\| = \sqrt{V_\alpha{}' V_\alpha} = \sqrt{(XU_\alpha)'(XU_\alpha)} = \sqrt{U_\alpha{}' X'XU_\alpha} = \sqrt{\lambda_\alpha} \neq 1$$

وحتى يكون هذا الشعاع الذاتي وحدوي يجب تقسيم مركبات هذا الشعاع على الطويلة، يكون:

$$V_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} XU_\alpha \quad ; \quad V_\alpha{}' V_\alpha = 1$$

و تمثل هذه العبارة أول عبارات الانتقال، فهي تمكننا من حساب الشعاع الذاتي الوحدوي V_α انطلاقاً من الشعاع الذاتي الوحدوي U_α. كما انه يمكننا تبسيط العبارة السابقة على النحو التالي:

$$V_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} XU_\alpha \Rightarrow V_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} F_\alpha \Rightarrow F_\alpha = \sqrt{\lambda_\alpha}\ V_\alpha$$

حيث أن F_α تمثل إحداثيات كل الأفراد على المحور α.

2.4. عبارة ثاني انتقال من $\Re^n$ نحو $\Re^p$

إذا قمنا بالتحليل في $\Re^n$ أي أننا نملك القيم الذاتية λ_α و الأشعة الذاتية V_α و نرغب الآن في الانتقال إلى $\Re^p$ بغيت دراسة و تحليل الأفراد، أي يجب علينا تحليل المصفوفة $(X'X)$ و استخراج القيم الذاتية λ_α و الأشعة الذاتية U_α.

غير أننا من التحليل السابق وجدنا أن الشعاع $(X'V_\alpha)$ هو شعاع ذاتي للمصفوفة $(X'X)$ و المرفق بالقيمة الذاتية λ_α ، يمكننا أن نضع مرحلياً: $U_\alpha = X'V_\alpha$

و حتى يكون الشعاع الذاتي أعلاه شعاع توجيه للمحور α يجب أن تكون طويلة هذا الشعاع تساوي الواحد، غير أن طويلة هذا الشعاع هي:

$$\|U_\alpha\| = \sqrt{U_\alpha{}' U_\alpha} = \sqrt{(X'V_\alpha)'(X'V_\alpha)} = \sqrt{V_\alpha{}' XX'V_\alpha} = \sqrt{\lambda_\alpha} \neq 1$$

وحتى يكون هذا الشعاع الذاتي وحدوي يجب تقسيم مركبات هذا الشعاع على الطويلة، يكون:

$$U_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} X'V_\alpha \quad ; \quad U_\alpha{}' U_\alpha = 1$$

و تمثل هذه العبارة ثاني عبارات الانتقال، فهي تمكننا من حساب الشعاع الذاتي الوحدوي U_α انطلاقاً من الشعاع الذاتي الوحدوي V_α. كما انه يمكننا تبسيط العبارة السابقة على النحو التالي:

$$U_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} X'V_\alpha \Rightarrow U_\alpha = \frac{1}{\sqrt{\lambda_\alpha}} G_\alpha \Rightarrow G_\alpha = \sqrt{\lambda_\alpha}\; U_\alpha$$

حيث أن G_α تمثل إحداثيات كل المتغيرات على المحور α.

5. عبارة إعادة تشكيل البيانات

إذا كنا نملك القيم الذاتية λ_α و الأشعة الذاتية U_α و V_α، و نرغب في إعادة تشكيل عناصر المصفوفة X فان عبارة إعادة تشكيل البيانات تعطى على النحو التالي:

$$X = \sum_{\alpha=1}^{p} \sqrt{\lambda_\alpha} V_\alpha U_\alpha{}'$$