

TRƯỜNG ĐẠI HỌC BÁCH KHOA TP. HCM  
Khoa KTXD - Bộ môn KTTNN

# PHÂN TÍCH DỮ LIỆU



Giảng viên: PGS. TS. NGUYỄN THÔNG  
E-mail: [nguyenthong@hcmut.edu.vn](mailto:nguyenthong@hcmut.edu.vn) or [nthong56@yahoo.fr](mailto:nthong56@yahoo.fr)  
Web: <http://www4.hcmut.edu.vn/~nguyenthong/>  
Tél. (08) 38 691 592- 098 99 66 719

PHÂN TÍCH DỮ LIỆU  
Chương 6: Phân tích thành phần chính (PCA)

**NỘI DUNG MÔN HỌC**

Chương 1. Thống kê mô tả (ôn).  
Chương 1bis. Xác suất & phân phối thống kê (ôn)..  
Chương 2. Khoảng tin cậy.  
Chương 3. Kiểm định thống kê.  
Chương 4. Phân loại dữ liệu (Classification).  
Chương 5. Phân nhóm dữ liệu (Cluster).  
Chương 6. **Phân tích thành phần chính (PCA).**  
Chương 7. Phân tích chuỗi thời gian.  
Chương 8. Hồi quy tuyến tính.  
Chương 9. Xử lý số liệu thực nghiệm.  
Chương 10. Giới thiệu phần mềm SPSS or R

PHÂN TÍCH DỮ LIỆU  
Chương 6: Phân tích thành phần chính (PCA)

**KIỂM TRA**

- Kiểm tra giữa kỳ 60 ph.
- Thi viết cuối môn học 90ph.
- Tiểu luận.

**Chú ý:**

- Cho phép mang tài liệu tham khảo.
- Mang theo máy tính làm bài tập.

8/29/2016 PGS. Dr. Nguyễn Thông 3

PHÂN TÍCH DỮ LIỆU  
Chương 6: Phân tích thành phần chính (PCA)

**NỘI DUNG**

- Giới thiệu.
- Khái niệm cơ bản.
- Nội dung phương pháp phân tích nhân tố chính.
- Ví dụ áp dụng.

8/29/2016 PGS. Dr. Nguyễn Thông 4

PHÂN TÍCH DỮ LIỆU  
Chương 6: Phân tích thành phần chính (PCA)

**GIỚI THIỆU**

Phân tích nhân tố các thành phần chính chuẩn hóa (PTNT) là một phương pháp phân tích dữ liệu (định tính) cho phép:

- Nhận biết xu thế chính của tập hợp dữ liệu khảo sát;
- Đánh giá và sắp hạng các cá nhân nghiên cứu dựa vào các đặc trưng (biến định lượng) của chúng.

8/29/2016 PGS. Dr. Nguyễn Thông 5

PHÂN TÍCH DỮ LIỆU  
Chương 6: Phân tích thành phần chính (PCA)

**GIỚI THIỆU**

- Đây là một phương pháp phân tích định tính dựa trên cơ sở giá trị các biến định lượng liên kết với từng cá nhân (hoặc khảo sát).
- Mỗi biến sẽ được đánh giá theo một thang điểm thích hợp.

8/29/2016 PGS. Dr. Nguyễn Thông 6

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÁI NIỆM CƠ BẢN**

**Momen quán tính chất điểm m đối với tâm O**

$$M_{m/o} = md^2$$

Chất  
điểm khối  
lượng m

8/29/2016 PGS. Dr. Nguyễn Thống 7

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÁI NIỆM CƠ BẢN**

**Momen quán tính chất điểm m đối với đường thẳng D**

$$M_{m/(D)} = md^2$$

Chất  
điểm khối  
lượng m

8/29/2016 PGS. Dr. Nguyễn Thống 8

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÁI NIỆM CƠ BẢN**

**Momen quán tính hệ thống chất điểm  $m_i$  đối với tâm O**

$$M_{m_i/o} = \sum m_i d_i^2$$

Chất  
điểm khối  
lượng  $m_i$

8/29/2016 PGS. Dr. Nguyễn Thống 9

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÁI NIỆM CƠ BẢN**

**TRONG TÂM G HỆ CHẤT ĐIỂM**

$$M_{m_i/G} = \sum m_i d_i^2 \Rightarrow \text{Min}$$

Chất  
điểm khối  
lượng  $m_i$

8/29/2016 PGS. Dr. Nguyễn Thống 10

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÁI NIỆM CƠ BẢN**

**Momen quán tính hệ thống chất điểm  $m_i$  đối với đường thẳng D**

$$M_{m_i/(D)} = \sum m_i d_i^2$$

Chất  
điểm khối  
lượng  $m_i$

8/29/2016 PGS. Dr. Nguyễn Thống 11

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÁI NIỆM KHÔNG GIAN p CHIỀU**

**KHÔNG GIAN 2D**

**KHÔNG GIAN 3D**

**KHÔNG GIAN p CHIỀU**

8/29/2016 PGS. Dr. Nguyễn Thống 12

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TÍCH VÔ HƯỚNG 2 VECTO**

Xét 2 vectơ  $u$  và  $v$  hợp với nhau bởi 1 góc  $\theta$ . Tích vô hướng của 2 vectơ như sau :

$$X \equiv \vec{u} \cdot \vec{v} = |\vec{u}| |\vec{v}| \cos \theta$$

**Chú ý :**

- Nếu thẳng góc ta có :  $X=0$
- Nếu ta có  $\vec{u} // \vec{v} \rightarrow \vec{u} \cdot \vec{v} = |\vec{u}| |\vec{v}|$

8/29/2016 PGS. Dr. Nguyễn Thống 13

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TÍCH VÔ HƯỚNG 2 VECTO**

Giả sử trong không gian 2D,  $u(\alpha_1, \beta_1)$  &  $v(\alpha_2, \beta_2)$

$$\vec{u} \cdot \vec{v} = \alpha_1 \alpha_2 + \beta_1 \beta_2$$

**Chú ý:**

$$|\vec{u}| = \sqrt{\alpha_1^2 + \beta_1^2}$$

8/29/2016 PGS. Dr. Nguyễn Thống 14

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TÍCH VÔ HƯỚNG 2 VECTO**

$OM \cdot \vec{U} = \alpha_1 x_1 + \alpha_2 x_2$

Nếu  $U$  là vectơ đơn vị  $\rightarrow \alpha_1 x_1 + \alpha_2 x_2 = OG$   
 $G$  là hình chiếu của  $M$  xuống  $\Delta$

8/29/2016 PGS. Dr. Nguyễn Thống 15

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**NHẬN XÉT**

- $\rightarrow$  Khi góc hợp bởi vectơ  $OM$  và vectơ  $U$  càng bé  $\rightarrow$  hình chiếu của  $OM$  xuống vectơ  $U$  KHÔNG BỊ « BIẾN DẠNG » nhiều (mất thông tin) so với  $OM$  ban đầu.
- $\rightarrow \alpha \rightarrow$  bé  $\rightarrow$  chất lượng hình chiếu của  $OM$  xuống trục chứa vectơ  $U$  càng tốt.

8/29/2016 PGS. Dr. Nguyễn Thống 16

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TÍCH VÔ HƯỚNG 2 VECTO**

Trong không gian  $p$  chiều

$$OM \cdot \vec{U} = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_p x_p$$

8/29/2016 PGS. Dr. Nguyễn Thống 17

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**CHÚ Ý**

Xét 2 vectơ  $u$  và  $v$  là vectơ đơn vị và hợp với nhau bởi 1 góc  $\theta$ . Tích vô hướng của 2 vectơ như sau :

$$X \equiv \vec{u} \cdot \vec{v} = \cos \theta$$

Tổng quát hoá trong không gian  $p$  chiều với  $u(x_1, x_2, \dots, x_p)$  và  $v(y_1, y_2, \dots, y_p)$  là các vectơ đơn vị  $\rightarrow$

$$\cos \theta = x_1 y_1 + x_2 y_2 + \dots + x_p y_p$$

8/29/2016 PGS. Dr. Nguyễn Thống 18

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**NHẬN XÉT**

Một cách tổng quát, trong không gian p chiều, xét giá trị tích vô hướng X của 2 vectơ đơn vị:

$X \rightarrow 0$ .  $\rightarrow$  2 vectơ gần trực giao nhau  
 $X \rightarrow 1$ .  $\rightarrow$  2 vectơ cùng chiều nhau  
 $X \rightarrow -1$ .  $\rightarrow$  2 vectơ ngược chiều nhau

8/29/2016 PGS. Dr. Nguyễn Thống 19

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**BÀI TẬP**

**Bài 1:** Tính góc hợp bởi các vectơ sau:  
 $U(3,4)$  &  $V(-2,5^{0.5})$   
 $U(1,3)$  &  $V(-6,2)$   
 $U(1,3,4,-2)$  &  $V(2,1,1,4)$

**Bài 2:** Tính giá trị hình chiếu của vectơ M lên vectơ U:  
 $M(1,5)$  &  $U(3^{0.5}/2,1/2)$   
 $M(2,-1,3,2)$  &  $U(1/2,-1/2,-1/2,1/2)$

8/29/2016 PGS. Dr. Nguyễn Thống 20

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**BÀI TẬP**

**Bài 3:** Tìm góc hợp bởi vectơ  $OM(3,4,3,2^{0.5})$  với các trục tọa độ trong không gian 4D.

**Bài 4:** Chứng minh trong không gian p chiều, gọi  $\alpha_i$  là góc hợp bởi một vectơ OM bất kỳ với các trục tọa độ  $\rightarrow \alpha_i$  có tính chất sau:

$$\sum_{i=1}^p \cos^2 \alpha_i = 1$$

8/29/2016 PGS. Dr. Nguyễn Thống 21

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KIẾN THỨC THÔNG KÊ CẦN THIẾT SỬ DỤNG TRONG PHÂN TÍCH NHÂN TỐ**

- Biến thông kê định lượng, các phân phối thông kê;
- Giá trị trung bình, phương sai, độ lệch chuẩn của một biến thông kê;
- Biến được trung tâm hóa và chuẩn hóa;

8/29/2016 PGS. Dr. Nguyễn Thống 22

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

- Hệ số tương quan tuyến tính của 2 biến.

$\rightarrow$  Đây là các kiến thức cần thiết không thể thiếu để tiếp tục nghiên cứu phương pháp này .

8/29/2016 PGS. Dr. Nguyễn Thống 23

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TRONG TÂM G**

Xét  $X(x_1, x_2, \dots, x_n) \rightarrow$

$$X_G = \frac{\sum x_i}{N}$$

The diagram shows a horizontal axis labeled X. On the left is the origin O. Several points are marked on the axis:  $x_{N-1}$ ,  $x_i$ ,  $x_1$ , and  $x_N$ . A point  $X_G$  is marked between  $x_i$  and  $x_1$ , representing the center of gravity.

8/29/2016 PGS. Dr. Nguyễn Thống 24

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Momen quán tính của n phần tử có khối lượng đơn vị so với trọng tâm G được xác định như sau :**

$$M_{\Sigma/G} = \sum_i 1 \cdot (x_i - x_G)^2$$

8/29/2016 PGS. Dr. Nguyễn Thống 25

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PHƯƠNG SAI V(X)**

**Xét  $X(x_1, x_2, \dots, x_n) \rightarrow$**

$$\sigma^2 = V(X) = \frac{\sum_i (x_i - x_G)^2}{N-1}$$

$\rightarrow$  **Mô tả sự phân tán của X chung quanh giá trị trung bình của các giá trị biến nghiên cứu**

8/29/2016 PGS. Dr. Nguyễn Thống 26

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**MỘT SỐ BIẾN ĐỔI BIẾN THỐNG KÊ**

8/29/2016 PGS. Dr. Nguyễn Thống 27

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**BIẾN TRUNG TÂM HÓA VÀ CHUẨN HÓA**

**Standardized**

Gọi X là một biến thống kê. Một biến được xem là trung tâm hóa và chuẩn hóa ký hiệu  $t_i$  suy từ  $x_i$  được định nghĩa:

$$t_i = \frac{x_i - \bar{X}}{\sigma}$$

**Tính chất:**  
 $t_i \rightarrow$  số **KHÔNG** đơn vị

$$\sum_i t_i = 0; \sigma_{t_i} = 1$$

8/29/2016 PGS. Dr. Nguyễn Thống 28

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Normalized**

$$t_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

**Tính chất:**  $\rightarrow t_i \rightarrow [0 \rightarrow 1]$   
 $\rightarrow t_i$  : không đơn vị

8/29/2016 PGS. Dr. Nguyễn Thống 29

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Adjusted Normalized**

$$t_i = 2 * \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} - 1$$

**Tính chất:**  $\rightarrow t_i \rightarrow [-1 \rightarrow 1]$   
 $\rightarrow t_i$  : không đơn vị

8/29/2016 PGS. Dr. Nguyễn Thống 30

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HIỆP PHƯƠNG SAI Cov(X,Y)**

$$\text{Cov}(X, Y) = \frac{\sum_{i=1}^N (x_i - \bar{X})(y_i - \bar{Y})}{N}$$

Hệ số **tương quan tuyến tính**  $r$  của 2 biến X và Y như sau :

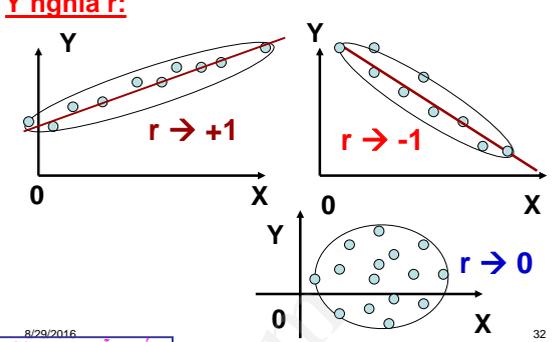
$$r = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y} \in [-1; +1]$$

**Chú ý:** Nếu biến đã chuẩn hoá  $\rightarrow \text{Cov}=r$

8/29/2016 PGS. Dr. Nguyễn Thống 31

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý nghĩa r:**



8/29/2016 PGS. Dr. Nguyễn Thống 32

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**BÀI TẬP**

**Bài 1:** Tính Cov của các vector sau:  
 $U(2,4,4,3,5)$  &  $V(1,3,5,3,7)$   
**Đáp số:**  $\text{Cov}(U,V)=1,92$

**Bài 2:** Tính r của các vector sau:  
 $U(3,2,5,5,1)$  &  $V(3,1,7,6,1)$   
**Đáp số:**  $\text{Cov}(U,V)=3,88$ ,  $\sigma_U=1,79$  &  $\sigma_V=2,79$   
 $\rightarrow r=0,78$

8/29/2016 PGS. Dr. Nguyễn Thống 33

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**BÀI TẬP**

**Bài 3:** Xét 2 vector  $X_1, X_2$  trong không gian 4D như sau:

	a	b	c	d
$X_1$	2	5	8	1
$X_2$	3	2	1	6

**Kiểm tra tính chất  $\rightarrow$  phương sai của biến  $X_i$  đã được chuẩn & trung tâm hoá bằng 1 ( $\sigma^2=V=1$ ).**

8/29/2016 PGS. Dr. Nguyễn Thống 34

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TRỤC QUẢN TÍNH**

Xét một cô thể (tập hợp vô số các phần tử có khối lượng 1 đơn vị) có trọng tâm là G.

- $\rightarrow$  Trục quán tính cấp 1 là đường thẳng đi qua G và hình chiếu của tất cả các chất điểm xuống trục này có PHƯƠNG SAI CỰC ĐẠI.
- $\rightarrow$  Trục quán tính cấp 2 là đường thẳng qua G, thẳng góc với trục 1 và hình chiếu tất cả các chất điểm xuống trục này có PHƯƠNG SAI CỰC ĐẠI.

8/29/2016 PGS. Dr. Nguyễn Thống 35

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Bài tập 0:** Xét 3 chất điểm có khối lượng 1 đơn vị ở các vị trí trong không gian 2D là  $A(-3,-2)$ ,  $B(2,5)$ ,  $C(1,-3)$ . Xác định trục chính 1 & trục chính 2.

8/29/2016 PGS. Dr. Nguyễn Thống 36

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Bài tập 1:** Xét tam giác cân với các đỉnh A(0,9), B(2,0) & C(-2,0). Xác định trục quán tính cấp 1 và cấp 2.

**Bài tập 2:** Xét tám gỗ đồng chất hình ellipse có phương trình:

$$\frac{(X-1)^2}{25} + \frac{(Y-2)^2}{4} = 1$$

Xác định trục quán tính chính cấp 1 và cấp 2

8/29/2016 PGS. Dr. Nguyễn Thống 37

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**LÝ THUYẾT**  
**PHƯƠNG PHÁP PHÂN TÍCH THÀNH PHẦN CHÍNH**  
**Principal Component Analysis (PCA)**

8/29/2016 PGS. Dr. Nguyễn Thống 38

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HÌNH THÀNH VẤN ĐỀ**

Khảo sát các yếu tố trong không gian p chiều (biến số, p>2).

→ Ta không thể phân tích số liệu một cách « trực quan ».

→ Vấn đề càng khó tưởng tượng khi p càng lớn.

→

8/29/2016 PGS. Dr. Nguyễn Thống 39

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HÌNH THÀNH VẤN ĐỀ**

**KHÔNG GIAN BIẾN**

8/29/2016 PGS. Dr. Nguyễn Thống 40

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**DẠNG SỐ LIỆU**

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TRONG PHÂN TÍCH NHÂN TỐ CHÍNH**  
**→ CÁC BIẾN ĐƯỢC TRUNG TÂM & CHUẨN HOÁ**

$$t_i = \frac{x_i - \bar{X}}{\sigma}$$

8/29/2016 PGS. Dr. Nguyễn Thống 42

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**DANG SỐ LIỆU PHÂN TÍCH**

	Var1	Var2	Var3
Obs1	0.472	-0.970	0.956
Obs2	-1.551	0.728	-1.036
Obs3	0.135	-1.213	0.159
Obs4	1.146	0.728	0.956
Obs5	-0.202	0.728	-1.036
Moy	0.00	0.00	0.00
Sigma	1.00	1.00	1.00

8/29/2016 PGS. Dr. Nguyễn Thống 43

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HÌNH THÀNH VẤN ĐỀ**

- Nhu cầu phân tích số liệu trong các mặt phẳng (không gian 2D).
- Chiều số liệu trong không gian p chiều xuống các mặt phẳng (không gian 2D).
- Chọn các mặt phẳng chiếu sao cho ÍT mất thông tin nhất do thực hiện phép chiếu.
- Đó chính là các **MẶT PHẪNG CHÍNH** trong phương pháp PTNT.

8/29/2016 PGS. Dr. Nguyễn Thống 44

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HÌNH THÀNH VẤN ĐỀ**

- Phương pháp PTNT chỉ chọn một số trục **ÍT NHẤT CÓ THỂ** để biểu diễn số liệu (thông thường từ 2 (2D) đến 4 (4D) trục).
- Đó là các trục **NHÂN TỐ CHÍNH** (Principal Component Axe).

8/29/2016 PGS. Dr. Nguyễn Thống 45

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Phân tích nhân tố → Giảm số biến ban đầu xuống còn ÍT biến hơn.**

→ Mỗi biến mới sẽ là **ĐẠI BIỂU** của một tập hợp biến cũ gọi là 1 **NHÂN TỐ** (item).

8/29/2016 PGS. Dr. Nguyễn Thống 46

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PRINCIPALE COMPONENT**

8/29/2016 PGS. Dr. Nguyễn Thống 47

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

→ Trục nhân tố được xác định theo phương pháp **PHƯƠNG SAI CỦA CÁC HÌNH CHIẾU BIẾN BAN ĐẦU XUỐNG TRỤC NHÂN TỐ CHÍNH LÀ CỰC ĐẠI.**

8/29/2016 PGS. Dr. Nguyễn Thống 48

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

→ Gọi  $\Delta_1, \Delta_2, \Delta_3, \Delta$  Là các thành phần CHÍNH 1, 2, 3, .....

**Tính chất:**

Variance các BIẾN TRÊN  $\Delta_1 >$   
 Variance các BIẾN TRÊN  $\Delta_2 >$   
 Variance các BIẾN TRÊN  $\Delta_3 >$   
 .....

8/29/2016 PGS. Dr. Nguyễn Thống 49

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PHƯƠNG TRÌNH XÁC ĐỊNH TRỰC CHÍNH  $\Delta_1, \Delta_2, \Delta_3, \Delta, \dots$**

8/29/2016 PGS. Dr. Nguyễn Thống 50

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

→ Giả thiết xem xét trong không gian p biến & n quan sát (observation):

→ Trục chính  $\Delta_1(\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n)$  ?

→  $\alpha_i$  (?) → tọa độ chuẩn hoá của trục  $\Delta_1$  trong KHÔNG GIAN n.

→ 
$$\sum_{i=1}^n \alpha_i^2 = 1$$

8/29/2016 PGS. Dr. Nguyễn Thống 51

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PHƯƠNG TRÌNH XÁC ĐỊNH TRỰC CHÍNH  $\Delta_1$**

$$V_1 \equiv \sum_{i=1}^p (\text{Variable}(i) * \Delta_1)^2 \quad [1] \Rightarrow \text{Max}$$

**Chú ý:** Variable(i) \*  $\Delta_1$

→ Hình chiếu Var(i) xuống  $\Delta_1$

8/29/2016 PGS. Dr. Nguyễn Thống 52

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

→ [1] là hàm theo các tham số  $\alpha_i$

→ Đạt cực trị → Đạo hàm riêng phần theo  $\alpha_i$  và = 0

→ Nhận được hệ phương trình giải ra xác định  $\alpha_i$

8/29/2016 PGS. Dr. Nguyễn Thống 53

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**SỬ DỤNG SOLVER TRONG EXCEL**  
**(BÀI TOÁN QUY HOẠCH TUYẾN TÍNH)**

→ **HÀM MỤC TIÊU & CÁC RÀNG BUỘC NẾU CÓ**

8/29/2016 PGS. Dr. Nguyễn Thống 54

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Giá trị  $V_1$  gọi là PHƯƠNG SAI GIẢI THÍCH CỦA TRỤC CHÍNH  $\Delta_1$**

**→  $V_1$  là max so với  $V_i$  của các trục chính  $\Delta_i$  còn lại**

**→**

8/29/2016 PGS. Dr. Nguyễn Thống 55

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Phương sai giải thích của trục chính  $\Delta_1$ :**

**→**

$$\text{Variance}(\Delta_1) = \frac{V_1}{\sum_{i=1}^p V_i} (\%)$$

8/29/2016 PGS. Dr. Nguyễn Thống 56

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PHƯƠNG TRÌNH XÁC ĐỊNH TRỤC CHÍNH  $\Delta_2$**

$$V_2 \equiv \sum_{i=1}^p (\text{Var}(i) * \Delta_2)^2 \Rightarrow \text{Max}$$

**Điều kiện:**  $\Delta_1 * \Delta_2 = 0$

**→ Hai trục là trục giao nhau**

8/29/2016 PGS. Dr. Nguyễn Thống 57

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PHƯƠNG TRÌNH XÁC ĐỊNH TRỤC CHÍNH  $\Delta_3$**

$$V_3 \equiv \sum_{i=1}^p (\text{Var}(i) * \Delta_3)^2 \Rightarrow \text{Max}$$

**Điều kiện:**  $\Delta_1 * \Delta_3 = \Delta_2 * \Delta_3 = 0$

**→ Các trục là trục giao nhau**

8/29/2016 PGS. Dr. Nguyễn Thống 58

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**→ Tiếp tục cho các trục  $\Delta_i$  kế tiếp**

**→ CHÚ Ý: CÁC TRỤC CHÍNH TẠO THÀNH MỘT HỆ TỌA ĐỘ TRỰC GIAO → ĐIỀU KIỆN CẦN & ĐỦ ĐỂ LÀM HỆ QUY CHIỀU.**

**→ DỪNG TRONG HỒI QUY TUYẾN TÍNH**

8/29/2016 PGS. Dr. Nguyễn Thống 59

**PHƯƠNG PHÁP ĐỊNH LƯỢNG TRONG QUẢN LÝ**  
**Chương 7: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC NHÂN TỐ CHÍNH**

**→ Trong không gian p chiều, trục nhân tố chính số 1 chỉ xu thế CHÍNH của ĐÁM MÂY dữ liệu.**

**→ Các trục nhân tố chính kế tiếp (2,3,4,...,p) chỉ các xu thế có khuynh hướng YẾU dần của ĐÁM MÂY số liệu.**

**→ HÌNH CHIỀU của ĐÁM MÂY xuống trục chính số 1 sẽ MẤT THÔNG TIN về số liệu là ÍT NHẤT.**

8/29/2016 PGS. Dr. Nguyễn Thống 60

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỰC NHÂN TỔ CHÍNH**

→ **HÌNH CHIỀU** của ĐÁM MÂY xuống **MẶT PHẪNG CHÍNH** số 1 sẽ **MẤT THÔNG TIN** về số liệu là **ÍT NHẤT**.

→ Thông thường ta **KHÔNG THỂ** nghiên cứu trong không gian  $p$  ( $p > 2$ ) chiều. Ta chỉ có thể nghiên cứu số liệu trong các **MẶT PHẪNG** (không gian 2D) → chiều các số liệu xuống các mặt phẳng → chọn các **MẶT PHẪNG ÍT MẤT THÔNG TIN NHẤT** → đó là các **MẶT PHẪNG CHÍNH**.

8/29/2016 PGS. Dr. Nguyễn Thống 61

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**MẶT PHẪNG CHÍNH**

→ Mặt phẳng chính số 1 tạo bởi trục chính  $\Delta_1$  và  $\Delta_2$ .

→ Mặt phẳng chính số 2 tạo bởi trục chính  $\Delta_1$  và  $\Delta_3$ .

→ Mặt phẳng chính số 3 tạo bởi trục chính  $\Delta_2$  và  $\Delta_3$ .

→ Mặt phẳng chính số 4, 5,.....

8/29/2016 PGS. Dr. Nguyễn Thống 62

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Bài tập**

Xét tập số liệu sau. Dùng Solver trong Excel để xác định trục chính  $\Delta_1$  &  $\Delta_2$ .

Observation	VAR1	VAR2	VAR3
1	8.00	2.00	9.00
2	2.00	9.00	4.00
3	7.00	1.00	7.00
4	10.00	9.00	9.00
5	6.00	9.00	4.00

8/29/2016 PGS. Dr. Nguyễn Thống 63

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HD: Xem Excel → PCA-Tutorial**

**Số liệu chuẩn & trung tâm hoá**

	Var1	Var2	Var3
Obs1	0.472	-0.970	0.956
Obs2	-1.551	0.728	-1.036
Obs3	0.135	-1.213	0.159
Obs4	1.146	0.728	0.956
Obs5	-0.202	0.728	-1.036
Moy	0.00	0.00	0.00
Sigma	1.00	1.00	1.00

8/29/2016 PGS. Dr. Nguyễn Thống 64

**PHƯƠNG PHÁP ĐỊNH LƯỢNG TRONG QUẢN LÝ**  
**Chương 7: Phân tích thành phần chính (PCA)**

**KẾT QUẢ**

Các thành phần chỉ hướng của các trục  $\Delta_1$   $\Delta_2$

	$\Delta_1$	$\Delta_2$
	-0.463	-0.276
	0.672	-0.133
	-0.247	-0.553
	-0.350	0.743
	0.388	0.219

8/29/2016 PGS. Dr. Nguyễn Thống 65

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**SỐ LIỆU BAN ĐẦU**

Phương pháp phân tích nhân tố các thành phần chính chuẩn hóa được áp dụng để tính toán một bảng số liệu 2D(n,p).

→ Chiều (hàng) biểu thị  $n$  **đối tượng** ( cá nhân, khảo sát).

→ Chiều còn lại (cột) biểu thị  $p$  **biến định lượng**  $x_i$  liên quan các **tính chất** của đối tượng.

8/29/2016 PGS. Dr. Nguyễn Thống 66

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**SỐ LIỆU BAN ĐẦU**

Biến Đối tượng	X <sub>1</sub>	X <sub>2</sub>	.	X <sub>j</sub>	X <sub>p</sub>
1	.	.	.	.	.
2	.	.	.	.	.
i	.	.	.	<b>X<sub>i,j</sub></b>	.
.	.	.	.	.	.
n	.	.	.	.	.

8/29/2016 PGS. Dr. Nguyễn Thống 67

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

với  $x_{ij}$  biểu thị giá trị của biến định lượng  $X_j$  của đối tượng (quan sát) thứ  $i$ .

- **Cột**  $X_i(a,b,c,...)$  → **vectơ biến** (biến thống kê, có  $n$  thành phần)
- **Hàng**  $N(x_1,x_2,...,x_p)$  → **định vị tọa độ CÁ NHÂN** (quan sát) trong **KHÔNG GIAN BIẾN**.
- Xét tính tương quan giữa các biến  $X_i$   
→ chuẩn hóa các biến ( $\sigma=1$ ) → dùng định nghĩa  $r(X_i, X_j)$

8/29/2016 PGS. Dr. Nguyễn Thống 68

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

- $r(..) \rightarrow +1$  → tương quan tuyến tính thuận chặt chẽ.
- $r(..) \rightarrow -1$  → tương quan tuyến tính nghịch chặt chẽ.
- $r(..) \rightarrow 0$  → **KHÔNG** tương quan.

8/29/2016 PGS. Dr. Nguyễn Thống 69

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**LOẠI BÀI TOÁN**

- **Loại 1:** Cột 1 là các Observation  
→ Bài toán loại tìm Principal Component &  
→ Có thể dùng kết quả trong nghiên cứu Hồi quy tuyến tính với các Principal Component là hệ **TỌA ĐỘ** tham chiếu.

8/29/2016 PGS. Dr. Nguyễn Thống 70

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ví dụ.**

Observation	VAR1	VAR2	VAR3
1 Obs1	8.00	2.00	9.00
2 Obs2	2.00	9.00	4.00
3 Obs3	7.00	1.00	7.00
4 Obs4	10.00	9.00	9.00
5 Obs5	6.00	9.00	4.00

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**LOẠI BÀI TOÁN**

- **Loại 2:** Cột 1 là các **CÁ NHÂN** → Bài toán loại tìm Principal Component & **SẮP HẠNG CÁ NHÂN**  
→ Có thể dùng kết quả trong nghiên cứu Hồi quy tuyến tính với các Principal Component là hệ **TỌA ĐỘ** tham chiếu.

8/29/2016 PGS. Dr. Nguyễn Thống 72

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ví dụ.**

	ten	coupdr	rev	serv	voler	retour
1	Arias	8.00	5.00	3.00	2.00	6.00
2	Becker	8.00	7.00	10.00	8.00	7.00
3	Borg	10.00	9.00	7.00	5.00	9.00
4	Carlsson	6.00	5.00	2.00	2.00	6.00
5	Cash	5.00	4.00	6.00	7.00	5.00
6	Clerc	7.00	6.00	4.00	3.00	5.00
7	Connors	7.00	10.00	4.00	6.00	10.00
8	Curren	4.00	4.00	8.00	7.00	6.00

**PHƯƠNG PHÁP ĐỊNH LƯỢNG TRONG QUẢN LÝ**  
**Chương 7: Phân tích thành phần chính (PCA)**

**KHÔNG GIAN BIẾN & KHÔNG GIAN CÁ NHÂN**

8/29/2016  
PGS. Dr. Nguyễn Thống 74

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÔNG GIAN BIẾN**

Xét bảng số liệu trên, với mỗi cá nhân ta có thể định nghĩa bằng 1 vectơ với các tọa độ  $(x_1, x_2, \dots, x_p)$ . Các trục  $x_i$  tạo thành 1 không gian  $p$  chiều, ta gọi là **KHÔNG GIAN BIẾN**.

→ Trong **KHÔNG GIAN** này sẽ **ĐỊNH VỊ** các **CÁ NHÂN**.

→ Khái niệm **ĐÁM MÂY SỐ LIỆU CÁ NHÂN**

8/29/2016  
PGS. Dr. Nguyễn Thống 75

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÔNG GIAN BIẾN  $X_{p=1,p}$**

8/29/2016  
PGS. Dr. Nguyễn Thống 76

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÔNG GIAN CÁ NHÂN**

8/29/2016  
PGS. Dr. Nguyễn Thống 77

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHÔNG GIAN CÁ NHÂN**

Xét bảng số liệu trên, với mỗi biến  $X$  ta có thể định nghĩa bằng 1 vectơ với các tọa độ  $(n_1, n_2, \dots, n_n)$ . Các trục  $n_i$  tạo thành 1 không gian  $n$  chiều, ta gọi là **KHÔNG GIAN CÁ NHÂN**.

→ Trong **KHÔNG GIAN** này sẽ **ĐỊNH VỊ** các **BIẾN  $X_i$** .

→ Khái niệm **ĐÁM MÂY SỐ LIỆU BIẾN**

8/29/2016  
PGS. Dr. Nguyễn Thống 78

**PHƯƠNG PHÁP ĐỊNH LƯỢNG TRONG QUẢN LÝ**  
**Chương 7: Phân tích thành phần chính (PCA)**

**KHÔNG GIAN CÁ NHÂN  $N_{N=1,N}$**

8/29/2016 PGS. Dr. Nguyễn Thống 79

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PHÂN TÍCH NHÂN TỐ CHÍNH VỚI SPSS**

8/29/2016 PGS. Dr. Nguyễn Thống 80

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**CỦA SỐ KHỞI ĐỘNG**

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHỞI ĐỘNG**

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**MỞ FILE CÓ SẴN**

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**DANG FILE SỐ LIỆU**

ten	coupr	rev	serv	voler	retour	smas
1 Arias	8.00	5.00	3.00	2.00	6.00	
2 Becker	8.00	7.00	10.00	8.00	7.00	
3 Borg	10.00	9.00				
4 Carlsson	5.00	5.00	2.00	2.00	0.00	
5 Cash	5.00	4.00	6.00	7.00	5.00	
6 Clerc	7.00	6.00	4.00	3.00	5.00	
7 Connors	7.00				10.00	

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Analyze =>Data Reduction=>Factor**

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**CHON BIẾN PHÂN TÍCH**

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**DESCRIPTIVES**

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**EIGENVALUE**

→ Giá trị liên kết cho từng trục chính.  
→ Được xác định theo **TỶ LỆ** của giá trị **PHƯƠNG SAI** hình chiếu đảm mây **BIẾN** xuống trục chính tương ứng.  
→ Tổng Engenvalue của p trục chính sẽ được hiệu chỉnh (từ giá trị phương sai) sao cho **BẢNG** số biến p (số trục chính tối đa).

8/29/2016 PGS. Dr. Nguyễn Thông 88

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**EXTRACTION**

**p/p PCA**  
**(Lưu ý khác với EFA - Phân tích khám phá - Principal factoring Analysis)**

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**ROTATION**

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**SCORES**

Factor Analysis: Factor Scores

- Save as variables
- Method:
  - Regression
  - Bartlett
  - Anderson-Rubin
- Display factor score coefficient matrix

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**OPTION**

Factor Analysis: Options

Missing Values:
 

- Exclude cases listwise
- Exclude cases pairwise
- Replace with mean

Coefficient Display Format:
 

- Sorted by size
- Suppress absolute values less than: 10

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA SỐ LIỆU KẾT QUẢ**

8/29/2016 93

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**THỐNG KÊ MÔ TẢ CÁC BIẾN**

Variable	Mean	Std. Deviation <sup>a</sup>	Analysis N <sup>a</sup>	Missing N
AMORTI	4,2857	2,1498	35	0
COUPDR	6,3429	1,4940	35	0
DOUBLE	4,2000	3,1230	35	0
JAMB	6,3714	1,4968	35	0
LOB	5,2286	1,7504	35	0
PASS	5,9714	1,5994	35	0
PHYSY	5,6571	1,5707	35	0
PSYCO	5,1429	1,9725	35	0
REGULA	5,5143	1,9459	35	0
RETOUR	5,7429	1,5213	35	0

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**MA TRẬN HỆ SỐ TƯƠNG QUAN TUYẾN TÍNH**

Correlation Matrix<sup>a</sup>

Correlation	AMORTI	COUPDR	DOUBLE	JAMB	LOB	PASS	PHYSY	PSYCO	R
AMORTI	1,000	-.004	,298	,231	,232	,216	,169	,157	
COUPDR	-.004	1,000	-.349	,533	,655	,767	,277	,412	
DOUBLE	,298	-.349	1,000	-.167	-.331	-.358	,182	-.105	
JAMB	,231	,533	-.167	1,000	,629	,688	,506	,579	
LOB	,232	,655	-.331	,629	1,000	,884	,265	,544	
PASS	,216	,767	-.358	,688	,884	1,000	,347	,626	
PHYSY	,169	,277	,182	,506	,265	,347	1,000	,595	
PSYCO	,157	,412	-.105	,579	,544	,626	,595	1,000	
REGULA	-.036	,661	,661	,661	,661	,661	,661	,661	
RETOUR	,086	,661	,661	,661	,661	,661	,661	,661	
REV	,237	,635	,635	,635	,635	,635	,635	,635	
SERV	,384	-.164	,646	,055	-.251	-.140	,505	,258	
SMASCH	,418	-.058	,563	,284	-.030	,033	,643	,530	
TOUCH	,904	,001	,377	,364	,162	,163	,236	,156	
YOLER	,519	-.288	,703	,116	-.243	-.221	,326	,152	

Giá trị nhỏ KHÔNG nằm chung trục & ngược lại

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

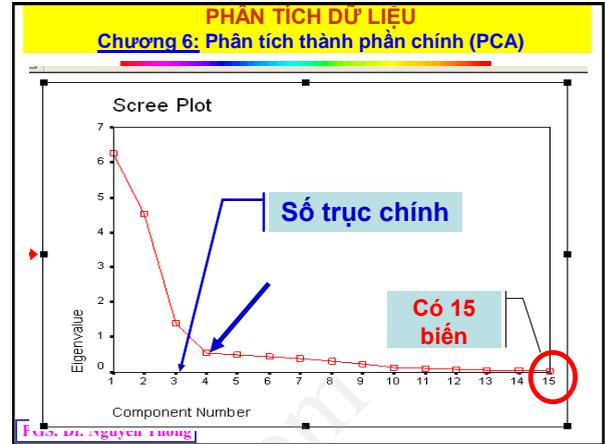
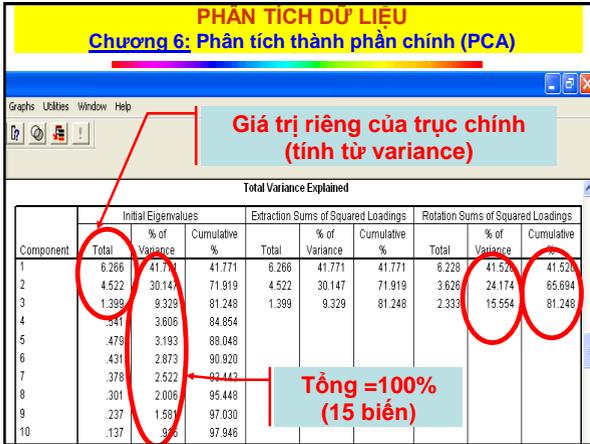
**KMO and Bartlett's Test**

> 0.7 → có ĐỦ chỉ tiêu cho mỗi factor

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.777
Bartlett's Test of Sphericity	Approx. Chi-Square	517.617
	df	105
		.000

< 0.05 → ma trận hệ số tương quan KHÔNG có TẤT CẢ h/số tương quan bằng 0 → Ok

PGS. Dr. Nguyễn Thông



**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HỆ SỐ TƯƠNG QUAN GIỮA THÀNH PHẦN CHÍNH & BIẾN**

→ **TRƯỜNG HỢP KHÔNG XOAY**

8/29/2016 PGS. Dr. Nguyễn Thống 99

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HỆ SỐ TƯƠNG QUAN GIỮA BIẾN VÀ TRỤC CHÍNH**

Trục chính CPAXe1

	Component Matrix <sup>a</sup>		
	1	2	3
AMORTI	-.399	.636	-.662
COUPDR	-.760	-.262	4.113E-03
DOUBLE	-.289	.739	8.30E-02
JAMB	.807	.143	4.113E-02
LOB	.853	-.194	.233
PASS	.941	-.162	.128
PHYSY	.533	.471	-.481
PSYCO	.764	.227	-.354
REGULA	.823	-.312	-.129
RETOUR	.897	-7.40E-02	-3.77E-02
REV	.904	2.984E-02	7.261E-02
SERV	1.00E-02	.893	-.267
SMASCH	.228	.883	-.312
TOUCH	.247	.748	.575
VOLEP	1.66E-02	.951	4.266E-02

8/29/2016 PGS. Dr. Nguyễn Thống 99

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HỆ SỐ TƯƠNG QUAN GIỮA THÀNH PHẦN CHÍNH & BIẾN**

→ **TRƯỜNG HỢP XOAY OPTION VARIMAX**

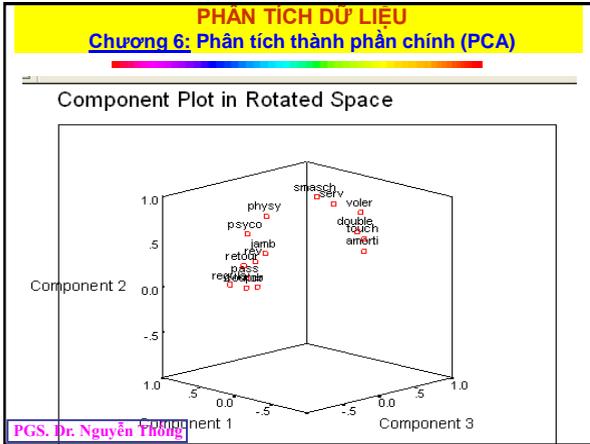
8/29/2016 PGS. Dr. Nguyễn Thống 101

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**H/số r<sup>2</sup> T/h có xoay**

	Rotated Component Matrix <sup>a</sup>		
	1	2	3
coupdr	.789	-.138	-.065
rev	.888	.084	.162
serv	-.108	.890	.254
voler	-.203	.758	.543
retour	.899	.057	.014
smasch	.105	.930	.231
jamb	.777	.185	.187
lob	.867	-.195	.173
amorti	.132	.197	.918
pass	.952	-.101	.110
regula	.862	-.100	-.197
touch	.126	.339	.906
psyco	.730	.465	-.104
physy	.470	.712	-.102
double	-.390	.629	.298

8/29/2016 PGS. Dr. Nguyễn Thống 101



**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**  
**TOA ĐỘ CÁ NHÂN TRÊN TRỤC CHÍNH**

Tennis1 - SPSS Data Editor

	ten	coupl	rev	fac1_1	fac2_1	fac3_1
1	Alias	8.00	5.00	.02880	-1.47959	-.62518
2	Becker	8.00	7.00	40431	1.94065	-.80924
3	Borg	10.00	9.00	2.61964	1.54676	-1.04101
4	Carlsson	6.00	5.00	12477	-1.22692	-1.12940
5	Cash	5.00	4.00	-.81524	1.16680	-1.09736
6	Clerc	7.00	6.00	-.00958	-1.09319	-.18186
7	Connors	7.00	10.00	2.18083	-.15125	.71479
8	Eden	4.00	4.00	-1.25490	.49957	-.10483

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**VÍ DỤ**

Để hiểu rõ phương pháp này chúng ta sẽ nghiên cứu thông qua một ví dụ cụ thể liên quan đến vấn đề phân tích và đánh giá các tay vợt chơi tennis hàng đầu trên thế giới dựa vào số liệu chấm điểm của tổ chức ATP (Association des joueurs de Tennis Professionnels).

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Dựa vào kết quả chơi tennis của các đấu thủ năm 1985, một Hội đồng gồm các chuyên gia trong lĩnh vực đã đánh giá cho điểm **35** tuyển thủ giỏi nhất, dựa vào **15** tiêu chuẩn khác nhau. Điểm đánh giá sẽ cho theo thang điểm 10.

PGS. Dr. Nguyễn Thống

- PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**
- Để phân tích ta dựa vào **15 tính chất (BIẾN)** sau đây :
- 1. Đánh bật trực tiếp (CPDT)
  - 2. Đánh rever (REV)
  - 3. Giao bóng (SERV)
  - 4. Đánh volé (VOLE)
  - 5. Trả bóng (RET)
  - 6. Đập bóng (SMASH)
- PGS. Dr. Nguyễn Thống

- PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**
- 7. Di chuyển (JBES)
  - 8. Tăng bóng (LOB)
  - 9. Giảm chấn (AMOR)
  - 10. Trả chéo sân (PASS)
  - 11. Đôi công (REGU)
  - 12. Bỏ nhỏ (TOUCH)
  - 13. Tâm lý (SPY)
  - 14. Thẻ lực (PHY)
  - 15. Đánh đôi (DBLE)
- PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Các tay vợt được đưa vào danh sách để đánh giá như sau :

1. Arias	2. Becker
3. Borg	4. Carlsson
5. Cash	6. Clerc
7. Connors	8. Curren
9. Edberg	10. Forget
11. Gerulaitis	12. Gomez
13. Gottfried	

8/29/2016 PGS. Dr. Nguyễn Thống 109

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

14. Higuera	15. Jaite
16. Jarryd	17. Krickstein
18. Kriek	19. Leconte
20. Lendl	21. Mac Enroe
22. Mayotte	23. Mercir
24. Noah	25. Nystroem

8/29/2016 PGS. Dr. Nguyễn Thống 110

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

26. Pecci	27. Pernfors
28. Sanchez	29. Solomm
30. Tanner	31. Teltscher
32. Tulasne	33. Vilas
34. Wilander	35. Zivojinovic

8/29/2016 PGS. Dr. Nguyễn Thống 111

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Lưu ý :**

- Chúng ta chỉ xét các trường hợp ở đó mỗi phân bố ở cột đều có phương sai khác không.

→ **Thật vậy nếu một biến đưa vào xét và có giá trị như nhau cho tất cả các đối tượng nghiên cứu thì nó không còn có giá trị cung cấp thông tin cho chúng ta.**

8/29/2016 PGS. Dr. Nguyễn Thống 112

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**MỘT SỐ KẾT QUẢ TỪ PHƯƠNG PHÁP PTNT**

8/29/2016 PGS. Dr. Nguyễn Thống 113

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

ten	ten	coupst	rev	serv	veler	retour	smash	jamb	ish	ame
1	Arias	8.00	5.00	3.00	2.00	6.00	3.00	6.00	5.00	
2	Becker	8.00	7.00	10.00	8.00	7.00	9.00	6.00	5.00	
3	Borg	10.00	9.00	7.00	5.00	9.00	9.00	10.00	7.00	
4	Carlsson	6.00	5.00	2.00	2.00	6.00	2.00	6.00	4.00	
5	Cash	5.00	4.00	6.00	7.00	5.00	7.00	6.00	3.00	
6	Clerc	7.00	6.00	4.00	3.00	5.00	3.00	6.00	6.00	
7	Connors	7.00	10.00	4.00	6.00	10.00	7.00	9.00	9.00	
8	Curren	4.00	4.00	8.00	7.00	6.00	6.00	5.00	3.00	
9	Edberg	5.00	8.00	8.00	8.00	6.00	6.00	6.00	3.00	
10	Forget	7.00	4.00	6.00	5.00	5.00	5.00	5.00	4.00	
11	Gerulaitis	6.00	5.00	6.00	10.00	4.00	6.00	8.00	4.00	
12	Gomez	8.00	5.00	4.00	5.00	4.00	5.00	5.00	5.00	
13	Gottfried	6.00	5.00							
14	Higuera	7.00	6.00							

8/29/2016 PGS. Dr. Nguyễn Thống 113

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KẾT QUẢ**

**CÁC THÔNG SỐ THỐNG KÊ CƠ BẢN**

Biến	Trung bình	Độ lệch chuẩn
CPDT	6.343	1.4725
REV	5.771	1.5508
SERV	5.429	2.2714
VOLE	5.343	2.4371
RET	5.743	1.4994
SMASH	5.571	1.9462

8/29/2016 PGS. Dr. Nguyễn Thống 115

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Biến	Trung bình	Độ lệch chuẩn
JBES	6.314	1.5264
LOB	5.229	1.7252
AMOR	4.286	2.1189
PASS	5.971	1.5764
REGU	5.514	1.9179
TOUCH	4.829	1.8744
PSY	5.143	1.9441
PHY	5.657	1.5481
DBLE	4.200	3.0780

8/29/2016 PGS. Dr. Nguyễn Thống 116

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**MA TRẬN TƯƠNG QUAN**

8/29/2016 PGS. Dr. Nguyễn Thống 117

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	CPDT	REV	SERV	VOLE	RET	SMASH	JBES	
CPDT	1.000							
REV	0.635	1.000						
SERV	-0.164	0.036	1.000					
VOLE	-0.288	-0.017	0.825	1.000				
RET	0.661	0.859	0.007	-0.085	1.000			
SMASH	-0.058	0.176	0.862	0.802	0.148	1.000		
JBES	0.550	0.706	0.027	0.071	0.734	0.267	1.000	
LOB	0.655	0.746	-0.251	-0.243	0.741	-0.030	0.635	1.000

**Bảng hệ số  $r=Cov(X,Y)$ , vì  $\sigma_x=\sigma_y=1$**

8/29/2016 PGS. Dr. Nguyễn Thống 118

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	JBES	LOB	AMOR	PASS	REGU	TOUCH	PSY	PHY	DBLE
AMOR	0.246	0.232	1.000						
PASS	0.692	0.864	0.216	1.000					
REGU	0.560	0.707	-0.036	0.818	1.000				
TOUCH	0.368	0.162	0.904	0.163	-0.095	1.000			
PSY	0.582	0.544	0.157	0.626	0.639	0.156	1.000		
PHY	0.481	0.265	0.169	0.347	0.329	0.236	0.595	1.000	
DBLE	-0.184	-0.331	0.298	-0.358	-0.390	0.377	-0.105	0.182	1.000

**Bảng hệ số  $r=Cov(X,Y)$ , vì  $\sigma_x=\sigma_y=1$**

8/29/2016 PGS. Dr. Nguyễn Thống 119

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	CPDT	REV	SERV	VOLE	RET	SMASH	JBES	LOB
AMOR	-0.004	0.237	0.384	0.579	0.086	0.418	0.246	0.232
PASS	0.767	0.862	-0.140	-0.221	0.855	0.033	0.692	0.864
REGU	0.676	0.693	-0.215	-0.380	0.702	-0.071	0.560	0.707
TOUCH	0.001	0.281	0.521	0.707	0.116	0.536	0.368	0.162

**Bảng hệ số  $r=Cov(X,Y)$ , vì  $\sigma_x=\sigma_y=1$**

8/29/2016 PGS. Dr. Nguyễn Thống 120

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**CPDT REV SERV VOLE RET SMASH JBES LOB  
AMOR PASS REGU TOUCH PSY PHY DBLE**

TOUCH	0.001	0.281	0.521	0.707	0.116	0.536	0.368
	0.162	0.904	0.163	-0.095	1.000		
PSY	0.412	0.608	0.258	0.152	0.640	0.530	0.582
	0.544	0.157	0.626	0.639	0.156	1.000	
PHY	0.277	0.455	0.505	0.326	0.356	0.643	0.481
	0.265	0.169	0.347	0.329	0.236	0.595	1.000
DBLE	-0.349	-0.200	0.646	0.703	-0.305	0.563	-0.184
	0.331	0.298	-0.358	-0.390	0.377	-0.105	0.182
							1.000

**Bảng hệ số  $r=Cov(X,Y)$ , vì  $\sigma_x = \sigma_y = 1$**

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**VECTƠ RIỀNG (phương sai trên trục chính)**

$\Sigma=15$  (số biến)

Trục I	Trục II	Trục III	Trục IV
6.2869	4.5158	1.4026	0.5396

Mức độ tham gia vào phương sai tổng (tỷ lệ phần trăm được giải thích bởi các trục chính).

Trục I	Trục II	Trục III	Trục IV
41.9%	30.1%	9.4%	3.6%

(41.9%=6.2869/15biến)

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HỆ SỐ CÁC BIẾN TRONG  
PHƯƠNG TRÌNH TUYẾN  
TÍNH GIỮA TRỤC CHÍNH  
& BIẾN**

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**QUAN HỆ TUYẾN GIỮA TRỤC  
CHÍNH  $\Delta$  VÀ CÁC BIẾN**

TRỤC CHÍNH  $\Delta_k = \sum_i \alpha_i X_i$

$k \rightarrow$  trục chính

Hệ số (xem bảng sau)

Biến đã trung tâm & chuẩn hoá ( $i=1 \rightarrow 15$ )

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TRƯỜNG HỢP CHON  
OPTION KHÔNG XOAY**

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Component Score Coefficient Matrix

	Component			
	1	2	3	4
coupdr	.121	-.058	.003	.659
rev	.144	.007	.052	.232
serv	.002	.197	-.191	.224
voler	-.011	.210	.035	.024
retour	.143	-.016	-.027	.159
smasch	.036	.195	-.223	-.139
jamb	.129	.032	.029	-.329
lob	.136	-.043	.167	.056
amorti	.038	.141	.473	-.208
pass	.150	-.036	.091	.184
regula	.131	-.069	-.092	.070
touch	.039	.165	.411	-.142
psyco	.122	.050	-.253	-.518
phys	.085	.104	-.344	-.198
double	-.046	.163	-.059	.862

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TRƯỜNG HỢP CHON**  
**OPTION VARIMAX**

8/29/2016 PGS. Dr. Nguyễn Thống 127

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Component Score Coefficient Matrix**

	Component		
	1	2	3
coupdr	.128	-.037	-.017
rev	.141	-.007	.061
serv	-.022	.268	-.055
voler	-.041	.155	.141
retour	.145	.017	-.018
smasch	.013	.287	-.079
jamb	.122	.024	.054
lob	.137	-.111	.130
amorti	.008	-.135	.476
pass	.152	-.063	.072
regula	.142	.007	-.102
touch	.007	-.081	.437
psyco	.119	.192	-.174
phy	.076	.283	-.225
double	-.068	.163	.033

8/29/2016 PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Hệ số trong phương trình tuyến tính giữa trục chính và các biến.**

	<u>Trục I</u>	<u>Trục II</u>	<u>Trục III</u>	<u>Trục IV</u>
CPDT	0.3040	-0.1197	0.0006	-0.4950
REV	0.3607	0.0185	-0.0569	-0.1685
SERV	0.0018	0.4202	0.2278	-0.1556
VOLE	-0.0314	0.4468	-0.0414	-0.0302
RET	0.3585	-0.0305	0.0349	-0.1164
SMASH	0.0872	0.4168	0.2645	0.1140

8/29/2016 PGS. Dr. Nguyễn Thống 129

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	<u>Trục I</u>	<u>Trục II</u>	<u>Trục III</u>	<u>Trục IV</u>
JBES	0.3262	0.0588	-0.0623	0.1924
LOB	0.3405	-0.0874	-0.1915	-0.0433
AMOR	0.0943	0.3009	-0.5561	0.1682
PASS	0.3765	-0.0715	-0.1024	-0.1248
REGU	0.3291	-0.1427	0.1141	-0.0331
TOUCH	0.0968	0.3532	-0.4852	0.0982
PSY	0.3033	0.1107	0.3029	0.4123
PHY	0.2093	0.2239	0.4082	0.1321
DBLE	-0.1181	0.3467	0.0703	-0.6276

8/29/2016 PGS. Dr. Nguyễn Thống 130

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TOA ĐỘ CÁ NHÂN (hoặc OBSERVATION) TRÊN TRỤC CHÍNH 1**

**$OBS(.)_{\Delta 1} = 0.304CPDT + 0.3607REV + \dots + 0.3585RET + \dots + 0.3805LOB + - 0.1181DOUBL$**

Tương tự cho các trục còn lại.

8/29/2016 PGS. Dr. Nguyễn Thống 131

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**CHÚ Ý**

**→ Hai kết quả trên cho từ 2 phần mềm khác nhau. Chúng sẽ hoàn toàn như nhau khi NHÂN VỚI 1 TỶ LỆ HẰNG SỐ cho CỘT**

8/29/2016 PGS. Dr. Nguyễn Thống 132

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**NGHIÊN CỨU BIẾN**

8/29/2016 133  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**NGHIÊN CỨU BIẾN**

**Cột 1 : Hệ số tương quan tuyến tính r giữa biến và các trục chính F.**

8/29/2016 134  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Component Matrix<sup>a</sup>**

	Component			
	1	2	3	4
pass	.941	-.162	.128	.100
rev	.904	.030	.073	.126
retour	.897	-.074	-.038	.086
lob	.853	-.194	.233	.030
regula	.823	-.312	-.129	.038
jamb	.807	.143	.041	-.178
psyco	.764	.227	-.354	-.280
coupdr	.760	-.262	.004	.356
physy	.533	.471	-.481	-.108
voler	-.067	.951	.050	.013
serv	.015	.893	-.267	.121
smasch	.228	.883	-.312	-.075
touch	.247	.748	.575	-.077
double	-.289	.739	-.083	.466
amorti	.239	.636	.662	-.113

8/29/2016 135  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Component	Matrix			r
	F1	F2	F3	
AMORTI	0.238523	0.635938	0.662062804	
COUPDR	0.759795	-0.26207	0.004128732	
DOUBLE	-0.28927	0.739312	-0.083047504	
JAMB	0.806536	0.142625	0.041126508	
LOB	0.853083	-0.19425	0.233027805	
PASS	0.94132	-0.16206	0.127576135	
PHYSY	0.532775	0.471409	-0.481203481	
PSYCO	0.76358	0.227476	-0.353774256	

8/29/2016 136  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Component	Matrix			r
	F1	F2	F3	
REGULA	0.82318	-0.3116	-0.129039797	
RETOUR	0.897319	-0.07397	-0.03769372	
REV	0.903516	0.029839	0.072613915	
SERV	0.014602	0.892555	-0.267075043	
SMASCH	0.227757	0.883421	-0.311706814	
TOUCH	0.246897	0.747919	0.57520454	
VOLER	-0.06658	0.950885	0.049657953	

8/29/2016 137  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**NGHIÊN CỨU CÁ NHÂN**

8/29/2016 138  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**NGHIÊN CỨU CÁ NHÂN**

**→ Tọa độ của cá nhân trên các trục chính**

8/29/2016 139

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

CÁ NHÂN	F1	F2	F3
Arias	0.029	-1.480	-0.625
Becker	0.404	1.941	-0.809
Borg	2.620	1.547	-1.041
Carlsson	-0.125	-1.227	-1.129
Cash	-0.815	1.166	-1.097
Clerc	-0.010	-1.093	-0.182
Connors	2.181	-0.151	0.715
Curren	-1.255	0.500	-0.105
Edberg	-0.316	1.120	0.043

8/29/2016 140

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

CÁ NHÂN	F1	F2	F3
Forget	-0.883	-0.221	0.031
Gerulati	-0.516	-0.212	2.839
Gomez	-0.635	-0.430	0.038
Gottfrie	-0.245	0.327	1.207
Higueras	0.421	-0.918	-1.034
Jaite	-0.476	-0.850	-1.026
Jarryd	-0.866	-0.065	0.074
Krickste	-0.357	-1.249	0.312
Kried	-0.825	0.040	0.386

8/29/2016 141

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

CÁ NHÂN	F1	F2	F3
Kried	-0.825	0.040	0.386
Leconte	-0.358	0.053	0.578
Lendl	1.684	0.507	-0.409
MacEnroe	1.149	1.191	2.515
Mayotte	-1.019	0.738	-0.454
Mercir	0.112	-1.253	2.083
Noah	-0.433	2.011	0.338
Nystroem	1.073	-0.179	-0.864
Pecci	-1.288	0.287	0.718
Pernfors	0.458	-0.933	0.903

8/29/2016 142

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

CÁ NHÂN	F1	F2	F3
Sanchez	-0.589	-0.596	-1.034
Solomm	1.134	-1.052	-0.715
Tanner	-0.671	1.869	-0.470
Teltsche	-0.322	-1.029	-0.233
Tulasne	-0.623	-0.830	-0.832
Vilas	1.461	-0.276	-0.415
Wilander	1.202	-0.203	0.401
Zivojino	-1.302	0.953	-0.707

8/29/2016 143

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PHÂN TÍCH & GIẢI THÍCH KẾT QUẢ**

8/29/2016 144

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**TRỤC CHÍNH**  
**(BIẾN COMPOSITE)**

8/29/2016 PGS. Dr. Nguyễn Thống 145

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**XÁC ĐỊNH CÁC BIẾN TƯƠNG QUAN**  
Từ kết quả ma trận hệ số tương quan ta có:

8/29/2016 PGS. Dr. Nguyễn Thống 146

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Component	Matrix		
	F1	F2	F3
AMORTI	0.238523	0.635938	<b>0.662062804</b>
COUPDR	0.759795	-0.26207	0.004128732
DOUBLE	-0.28927	0.739312	-0.083047504
JAMB	0.806536	0.142625	0.041126508
LOB	<b>0.853083</b>	-0.19425	0.233027805
PASS	<b>0.94132</b>	-0.16206	0.127576135
PHYSY	0.532775	0.471409	-0.481203481
PSYCO	0.76358	0.227476	-0.353774256

8/29/2016 PGS. Dr. Nguyễn Thống 147

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Component	Matrix		
	F1	F2	F3
REGULA	<b>0.82319</b>	-0.3116	-0.129039797
RETOUR	<b>0.897319</b>	-0.07397	-0.03769372
REV	<b>0.903516</b>	0.029839	0.072613915
SERV	0.014602	<b>0.892555</b>	-0.267075043
SMASCH	0.227757	<b>0.883421</b>	-0.311706814
TOUCH	0.246897	0.747919	<b>0.57520454</b>
VOLER	-0.06658	<b>0.950885</b>	0.049657953

8/29/2016 PGS. Dr. Nguyễn Thống 148

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**DÙNG BẢN KẾT QUẢ TƯƠNG QUAN GIỮA TRỤC CHÍNH & BIẾN**

**Trục I** : Pass, Rev, Retour, Lob  
**Trục II** : Vol, Serv, Smash  
**Trục III** : Touch, Amor

8/29/2016 PGS. Dr. Nguyễn Thống 149

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC CHÍNH**  
**(Biến composite)**

**Trục I** : Pass, Rev, Retour, Lob  
→ Thể hiện bản lĩnh “phòng thủ”  
→ Đây là trục đại biểu cho tiêu chí phòng thủ TỐT.

8/29/2016 PGS. Dr. Nguyễn Thống 150

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC CHÍNH (Biến composite)**

**Trục I :** Pass, Rev, Retour, Lob

8/29/2016 151

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC CHÍNH (Biến composite)**

**Trục I :**

Biến	Hệ số tương quan tuyến tính r với trục F <sub>1</sub>
PASS	0.944
REV	0.904
RET	0.898
LOB	0.853
REGU	0.825

8/29/2016 152

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC CHÍNH (Biến composite)**

**Trục II :** Vol, Serv, Smash

→ Thể hiện lối đánh “tấn công”

→ Đây là trục đại biểu cho tiêu chí đánh tấn công nhanh.

8/29/2016 153

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC CHÍNH (Biến composite)**

**Trục II :** Vol, Serv, Smash

8/29/2016 154

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC CHÍNH (Biến composite)**

**Trục II :**

Biến	Hệ số tương quan tuyến tính r với trục F <sub>2</sub>
VOL	0.949
SER	0.892
SMAS	0.885
TOU	0.75
AMO	0.639

8/29/2016 155

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC CHÍNH (Biến composite)**

**Trục III :** Touch, Amor

→ Lối đánh thiên về “khéo léo”, ngược với tính chất dùng thể lực (Phy).

8/29/2016 156

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ý NGHĨA TRỤC CHÍNH (Biên composite)**

**Trục III : Touch, Amor**

Phy Amor Touch  $\Delta_3$

8/29/2016 PGS. Dr. Nguyễn Thống 157

**PHƯƠNG PHÁP ĐỊNH LƯỢNG TRONG QUẢN LÝ**  
**Chương 7: Phân tích thành phần chính (PCA)**

**KHẢ NĂNG GIẢI THÍCH SỐ LIỆU GỐC CỦA CÁC TRỤC CHÍNH**

8/29/2016 PGS. Dr. Nguyễn Thống 158

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHẢ NĂNG CÁC TRỤC CHÍNH F**

→ Trục chính F1 giải thích được 41,77% **BẢN CHẤT SỐ LIỆU GỐC** của các **CÁ NHÂN**.  
 → Tiếp theo là trục F2 được 30,14%.  
 → Trục F3 được 9,4%.  
 → Trục F4 được 4,6% (có thể bỏ qua vì <math>1/15=6,6\%</math> giá trị TB mỗi trục).

**Với 3 trục chính đầu đã giải thích được 81,3% tính chất CÁ NHÂN (khá tốt).**

8/29/2016 PGS. Dr. Nguyễn Thống 159

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KHẢ NĂNG CÁC TRỤC CHÍNH F**

→ **Mặt phẳng nhân tố thứ 1 (F1, F2) chiếm 71,9% sự khác biệt tổng quát của 35 đấu thủ xét trên 15 tiêu chuẩn đưa ra.**  
 → **Mặt phẳng thứ 2 có thể giải thích 52,2% số liệu gốc.**

8/29/2016 PGS. Dr. Nguyễn Thống 160

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**NGHIÊN CỨU CÁ NHÂN (Nghiên cứu khẳng định)**

8/29/2016 PGS. Dr. Nguyễn Thống 161

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Cá nhân Tọa độ trên trục I**

<b>BOR</b>	<b>2.62</b>
<b>CON</b>	<b>2.18</b>
<b>WIL</b>	<b>1.20</b>
<b>LEN</b>	<b>1.68</b>
<b>MAC</b>	<b>1.15</b>

8/29/2016 PGS. Dr. Nguyễn Thống 162

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Cá nhân Toa độ trên trục II**

NOA	2.01
BEC	1.94
TAN	1.87
MAC	1.19
EDB	1.12

8/29/2016 PGS. Dr. Nguyễn Thống 163

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Cá nhân Toa độ trên trục III**

GER	2.84
MAC	2.15
MEC	2.08

8/29/2016 PGS. Dr. Nguyễn Thống 164

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**SẮP HẠNG CÁ NHÂN**

→ Khai thác số liệu từ kết quả cho bởi SPSS.

8/29/2016 PGS. Dr. Nguyễn Thống 165

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**PHƯƠNG TRÌNH TUYẾN TÍNH CỦA CÁ NHÂN BIỂU DIỄN THEO CÁC TRỤC CHÍNH**

8/29/2016 PGS. Dr. Nguyễn Thống 166

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**CÂU HỎI**

Từ số liệu cho bởi SPSS, hãy cho biết cách xác định  $\cos(\alpha)$  → **GIÁ TRỊ** xác định chất lượng trình bày của trục chính đối với cá nhân N.

8/29/2016 PGS. Dr. Nguyễn Thống 167

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**BAI TAP**

Anh (Chị) hãy hình thành các chủ đề mà ta có thể sử dụng phương pháp CPA để nghiên cứu:

- Hình thành bài toán,
- Chuẩn bị dạng số liệu.
- Các kết luận dự kiến rút ra từ kết quả với SPSS.

8/29/2016 PGS. Dr. Nguyễn Thống 168

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Bài tập: Một nghiên cứu đánh giá nhằm sắp hạng các nhà thầu xây dựng để phục vụ lựa chọn nhà thầu.**

→ Một nghiên cứu với 40 nhà đầu tư để đánh giá 11 tiêu chí (thang điểm 10) thể hiện chất lượng.

8/29/2016 PGS. Dr. Nguyễn Thống 169

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Số liệu thăm dò và kết quả cho bởi phương pháp PTNT như sau.**

a. Dùng ma trận tương quan xác định các biến có tương quan cao.

b. Xác định các trục chính và giải thích các kết quả.

8/29/2016 PGS. Dr. Nguyễn Thống 170

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

<b>BIẾN</b>	<b>TÊN</b>
• Safety reco.	Sar
• Value profess.	Vap
• Reput. On time	Ron
• Rept. Serv.	Res
• Capital	Cap
• Experience	Exp
• Post. Serv	Pos
• Effi. Organi.	Efo
• Exp. Similaire pro.	Esp
• Personal team.	Pet
• Depth of tech. resource	Dtr

8/29/2016 PGS. Dr. Nguyễn Thống 171

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

MỘT SỐ KẾT QUẢ

8/29/2016 PGS. Dr. Nguyễn Thống 172

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	Safety reco.	Value Pro.	Reput. On time	Rept. Serv.	Capital	Experie nce
1	9	5	3	2	6	3
2	8	7	10	8	7	9
3	10	9	7	5	9	9
4	6	5	2	2	6	2
5	6	4	6	7	5	7
6	10	6	4	3	5	3
7	6	10	9	6	7	7
8	6	4	8	7	6	6
9	6	8	8	8	6	6
10	9	4	6	5	5	5

8/29/2016 PGS. Dr. Nguyễn Thống 173

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	Post. Serv	Effi. Organi	Exp. Similaire pro.	Personal team.	Depth of tech. resource
1	6	5	2	6	6
2	6	5	3	6	5
3	10	7	4	10	10
4	6	4	2	6	6
5	6	3	3	4	4
6	6	6	3	6	6
7	9	9	6	9	8
8	5	3	4	4	4
9	6	3	4	5	5
10	5	4	4	5	4

8/29/2016 PGS. Dr. Nguyễn Thống 174

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	Safety reco.	Value. Pro.	Reput. On time	Rept. Serv.	Capital	Experi nce
11	5	5	6	10	4	6
12	6	5	4	5	4	5
13	5	5	6	9	6	7
14	8	6	4	2	6	3
15	5	5	3	4	5	4
16	6	5	5	5	4	5
17	10	5	3	3	5	4
18	6	5	5	6	5	6
19	6	6	7	7	6	6
20. Dr. Nguyễn Thống	8	7	5	8	6	6

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	Post. Serv	Effi. Organi	Exp. Similair e pro.	Personal team.	Depth of tech. resource
11	8	4	10	5	5
12	5	5	4	5	4
13	6	7	5	9	9
14	6	5	2	7	8
15	5	4	2	5	5
16	5	4	4	5	5
17	5	5	5	5	5
18	9	4	3	4	1
19	6	4	4	6	4
20. Dr. Nguyễn Thống	8	4	4	8	9

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**(Số liệu còn lại xem trong PTNT\_Tennis\_Beer...xls)**

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	Mean	Std. Deviation	Analysis N
CAP	5.55	2.601	40
EXP	6.00	1.679	40
DTR	6.18	1.781	40
EFO	6.68	1.639	40
ESP	5.33	1.831	40
PET	4.18	2.062	40
POS	5.65	2.020	40
RES	5.45	2.320	40
ROT	5.90	1.722	40
SAR	5.78	2.069	40
VAP	6.75	1.780	40

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**KẾT QUẢ XEM TRONG**  
**Output1\_ChonThau.ops**

**→ Phân tích kết quả**

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**BÀI TẬP**

**Bài 2.** Để có thể tìm ra một chiến lược cho phép thâm nhập thị trường tiêu thụ bia ở VN có hiệu quả, một Công ty nước ngoài đã thực hiện một cuộc thăm dò ý kiến người tiêu dùng về các loại bia đang kinh doanh ở thị trường, với các đặc trưng phân tích đánh giá cụ thể. Các loại bia sẽ được đánh giá theo thang điểm 10 với các đặc trưng nghiên cứu. Kết quả thống kê được trình bày ở bảng sau :

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**Ghi chú :**

**FOR : Hình dạng PUB : Quảng cáo**

**VOL : Thể tích ALC : Tỷ lệ alcohol**

**RES : Phân phối EFI : Tác dụng đầu tiên**

**PSY : Tâm lý PRI : Giá**

**PAR : Mùi vị EFS : Tác dụng lâu dài.**

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Tên	TIG	GUI	BGI	HEI	BTH	SMI	HAM	333
Tính chất								
FOR	9	9	6	10	4	7	5	9
PUB	7	4	5	5	3	10	10	5
VOL	5	6	5	5	2	8	8	6
ALC	9	10	4	10	6	7	4	10
RES	9	7	5	6	3	9	10	7
EFI	10	9	5	10	3	6	5	9
PSY	10	6	5	7	3	8	8	6
PRI	4	6	4	10	2	3	7	8
PAR	9	7	6	9	3	6	5	8
EFS	9	7	5	9	2	6	5	10

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**a - Các Anh (Chị) có thể nói gì về các trục chính F1, F2... và các mặt phẳng chính tương ứng ?**

**b - Biện nào được giới thiệu tốt cho mặt phẳng chính số 1.**

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**c - Loại bia nào đang “chiếm lĩnh” thị trường theo kết quả này . Theo Anh (Chị), chiến lược nào phải được chọn để thâm nhập thị trường. Hãy giải thích tác dụng yếu tố giá (PRI) trong trường hợp này .**

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**MỘT SỐ KẾT QUẢ**

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Tên	TIG	GUI	BGI	HEI	BTH	SMI	HAM	333
Tính chất								
FOR	0.74	0.74	-0.63	1.19	-1.53	-0.17	-1.08	0.74
PUB	0.33	-0.8	-0.43	-0.43	-1.18	1.47	1.47	-0.43
VOL	-0.33	0.2	-0.33	-0.33	-1.89	1.24	1.24	0.2
ALC	0.57	0.95	-1.34	0.95	-0.57	-0.19	-1.34	0.95
RES	0.86	0	-0.86	-0.43	-1.72	0.86	1.29	0
EFI	1.07	0.7	-0.79	1.07	-1.53	-0.42	-0.79	0.7
PSY	1.58	-0.29	-0.76	0.18	-1.7	0.64	0.64	-0.29
PRI	-0.55	0.18	-0.55	1.65	-1.28	-0.92	0.55	0.92
PAR	1.15	0.18	-0.3	1.15	-1.75	-0.3	-0.79	0.67
EFS	1.19	0.32	-0.54	1.19	-1.84	-0.11	-0.54	0.32

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Biến	Giá trị trung bình	Độ chênh lệch
FOR	7.375	2.0578
PUB	6.125	2.4717
VOL	5.625	1.7984
ALC	7.5	2.4495
RES	7	2.1794
EFI	7.125	2.5218
PSY	6.625	1.9961
PRI	5.5	2.5495
PAR	6.625	1.9325
EFS	6.25	2.1651

8/29/2016 187  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	FOR	PUB	VOL	ALC	RES	EFI	PSY	PRI	PAR	EFS
FOR	1									
PUB	-0.132	1								
VOL	0.173	0.826	1							
ALC	0.88	-0.341	-0.071	1						
RES	0.251	0.859	0.861	0.047	1					
EFI	0.979	-0.083	0.176	0.86	0.341	1				
PSY	0.46	0.719	0.622	0.217	0.891	0.556	1			
PRI	0.608	-0.03	0.286	0.46	0.202	0.612	0.209			
PAR	0.947	0.01	0.211	0.726	0.356	0.959	0.612	0.596	1	
EFS	0.933	0.134	0.313	0.707	0.477	0.956	0.716	0.589	0.978	1

8/29/2016 188  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Giá trị riêng : 5.6769    3.0544    0.714

Phần trăm tham gia tính trên tổng phương sai của hệ thống đối với từng trục :

56.8%            30.5%            7.1%

- Cột 1 : Tương quan giữa biến và các trục chính
- Cột 2 : Cột 1 bình phương

8/29/2016 189  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Biến	Thành phần chính					
	Trục 1		Trục 2		Trục 3	
FOR	0.9175	0.8418	-0.3650	0.1332	0.0307	0.0009
PUB	0.2351	0.0553	0.9539	0.9099	0.0125	0.0002
VOL	0.4559	0.2078	0.7750	0.6006	-0.2892	0.0836
ALC	0.7200	0.5184	-0.5507	0.3033	0.1067	0.0114
RES	0.5936	0.3524	0.7762	0.6026	0.0455	0.0021
EFI	0.9457	0.8943	-0.3064	0.0939	0.0726	0.0053
PSY	0.7470	0.5580	0.5597	0.3133	0.2887	0.0833
PRI	0.6421	0.4123	0.1902	0.0360	0.7140	0.5111

8/29/2016 190  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Cột 1 : Tọa độ của các cá thể nghiên cứu trên trục chính  
Cột 2 : Chất lượng giới thiệu (cos2)

Hiệu	Trục 1		Trục 2		Trục 3	
TIG	2.5156	0.6513	0.2606	0.0070	1.7179	0.3038
GUI	1.0358	0.3256	-1.1609	0.4090	-0.0769	0.0018
BGI	-2.1414	0.7886	-0.1933	0.0064	-0.0924	0.0015
HEI	2.5663	0.6108	-1.7600	0.2873	-0.8037	0.0599
BTH	-4.9172	0.8849	-1.6506	0.0997	0.4288	0.0067
SMI	0.1038	0.0016	2.4053	0.8419	0.5703	0.0473
HAM	-0.6543	0.0356	3.1909	0.8456	-1.0596	0.0932
333	1.4915	0.5355	-1.0921	0.2871	-0.6845	0.1128

8/29/2016 191  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Bài 2. Đề nghiên cứu về mức sống của 20 nước có kỹ nghệ phát triển nhất trên thế giới, người ta đã tiến hành một điều tra về các khoảng chi tiêu khác nhau trên đầu người. Kết quả số liệu điều tra được trình bày ở bảng sau.

**Ghi chú : ALIM : Thực phẩm; HABI : ăn mặc; LOGE : ở MEU : Nội thất; SAN : Thuốc men; TRA : Di chuyển; LOIS : Giải trí AUTR : Khác**

8/29/2016 192  
PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	Pay.1	ALI M	HAB I	LOG E	MEU	SAN	TRA	LOI S	AUT R
1	BEL	1548	587	1332	757	772	885	472	1027
2	DK	2093	531	2235	608	158	1551	856	870
3	RFA	1556	738	1743	767	1303	1332	816	899
4	GRE	1074	231	313	222	79	404	121	251
5	FRA	1643	580	1503	678	711	1314	577	1011
6	IRL	1722	255	495	260	102	518	365	272
7	ITA	1589	591	970	558	359	819	546	1056
8	LUI	1702	481	1534	672	488	1237	252	946
9	HOL	1392	533	1418	550	923	795	693	992
10	GB	1153	440	1237	113	81	992	589	1191

8/29/2016 PGS. Dr. Nguyễn Thống 193

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

	Pay.1	ALIM	HABI	LOGE	MEU	SAN	TRA	LOIS	AUTR
11	ESP	1045	277	584	272	132	533	258	736
12	POR	753	187	99	173	92	296	116	266
13	HOR	2302	724	1514	721	325	1555	772	943
14	SUE	1921	614	2085	530	204	1318	806	510
15	SUI	3479	583	2343	643	1115	1360	1214	1226
16	AUT	1644	766	1406	505	367	1154	428	1195
17	FIN	1951	420	1406	549	276	1314	753	903
18	USA	1589	767	2307	672	1642	1789	1079	1818
19	CAN	1452	516	1923	766	347	1286	891	1213
20	JAP	2056	594	1761	506	990	889	913	1672

8/29/2016 PGS. Dr. Nguyễn Thống 194

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

Áp dụng phương pháp phân tích nhân tố các thành phần chính chuẩn hóa, Anh (Chi) hãy :

a - Thực hiện các bước cơ bản để phân tích và giải thích kết quả.

b - Những nước nào được giới thiệu tốt nhất, xấu nhất bởi mặt phẳng chính (1,2). Hãy giải thích ý nghĩa vật lý.

8/29/2016 PGS. Dr. Nguyễn Thống 195

**PHÂN TÍCH DỮ LIỆU**  
**Chương 6: Phân tích thành phần chính (PCA)**

**HẾT CHƯƠNG**

**XIN CẢM ƠN**

8/29/2016 PGS. Dr. Nguyễn Thống 196