NOTAZIONE

$\underset{n\times p}{X}$ = MATRICE DEI DATI ORIGINARIA

$\underset{n\times p}{\tilde{X}}$ = MATRICE DEGLI SCOSTAMENTI DALLA MEDIA

$\underset{n\times 1}{\mathbf{1}}$ = VETTORE FORMATO DA ELEMENTI TUTTI UGUALI AD 1 $= \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

OSSERVAZIONE $\underset{1\times n}{\mathbf{1}^T} \underset{n\times p}{\tilde{X}} = \underset{1\times p}{0}$

$\underset{n\times p}{\tilde{X}} = \underset{n\times n}{H} \cdot \underset{n\times p}{X}$ DOVE $\underset{n\times n}{H} = I_n - \mathbf{1}\mathbf{1}^T/n$

H = MATRICE SIMMETRICA E IDEMPOTENTE

$H = H \cdot H$

$\underset{p\times p}{S}$ = MATRICE DI VARIANZE E COVARIANZE

$$\underset{p\times p}{S} = \frac{\tilde{X}^T \tilde{X}}{n-1} = \frac{X^T H X}{n-1}$$

$\underset{p\times p}{D}$ = MATRICE DIAGONALE CHE CONTIENE SULLA DIAGONALE PRINCIPALE GLI SCOSTAMENTI QUADRATICI MEDI (CAMPIONARI) DELLE VARIABILI ORIGINARIE $D = \mathrm{diag}(s_1, \dots, s_p)$

$\underset{n \times p}{Z}$ = MATRICE CHE CONTIENE GLI SCOSTAMENTI STANDARDIZZATI

$$Z = \underset{n \times p}{\tilde{X}} \underset{p \times p}{D^{-1}} = H X D^{-1}$$

$\underset{p \times p}{R}$ = MATRICE DI CORRELAZIONE

$$R = \frac{Z'Z}{n-1} = \frac{D^{-1} X^T H X D^{-1}}{n-1}$$

$\underset{n \times p}{Y}$ = MATRICE CHE CONTIENE LE COMPONENTI PRINCIPALI

$Y = \tilde{X} V$ MATRICE CHE CONTIENE GLI AUTOVETTORI DELLA MATRICE S

$Y = Z V$ MATRICE CHE CONTIENE GLI AUTOVETTORI DELLA MATRICE R

$$Y = (\underset{n \times 1}{Y_1}\ \underset{n \times 1}{Y_2} \ldots \underset{n \times 1}{Y_p})$$

VETTORE CHE CONTIENE LA PRIMA COMPONENTE PRINCIPALE (PC)

VETTORE CHE CONTIENE LA P-ESIMA COMPONENTE PRINCIPALE (PC)

$$\underset{n \times 1}{Y_i} = \underset{n \times p}{Z} \underset{p \times 1}{V_i}$$

i ESIMO AUTOVETTORE ASSOCIATO ALL'I-ESIMO AUTOVALORE DELLA MATRICE S

SCOMPOSIZIONE SPETTRALE

QUALSIASI MATRICE $\underset{p \times p}{S}$ SIMMETRICA PUO' ESSERE SCOMPOSTA COME SEGUE

$$\underset{p \times p}{S} = \underset{p \times p}{V} \underset{p \times p}{\Lambda} \underset{p \times p}{V^T}$$

$\underset{p \times p}{V}$ = MATRICE ORTONORMALE ($V^T V = I_p$) CHE CONTIENE GLI AUTOVETTORI DELLA MATRICE S

$\underset{p \times p}{\Lambda}$ = MATRICE DIAGONALE CHE CONTIENE SULLA DIAGONALE PRINCIPALE GLI AUTOVALORI DELLA MATRICE S

$$\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_p) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_p \end{pmatrix}$$

$$\underset{p \times p}{V} = (\underset{p \times 1}{v_1}, \underset{p \times 1}{v_2}, \dots, \underset{p \times 1}{v_p})$$

PRIMO AUTOVETTORE ASSOCIATO A λ_1

p-ESIMO AUTOVETTORE ASSOCIATO A λ_p

$$S = (v_1 \ v_2 \ \dots \ v_p) \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_p \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_p^T \end{pmatrix}$$

$$S = \sum_{i=1}^{p} \underset{p \times 1}{v_i} \underset{1 \times p}{v_i^T} \cdot \lambda_i$$

OBIETTIVO PRIMA COMPONENTE PRINCIPALE

DATA $\tilde{X}$ OPPURE Z SI CERCA IL VETTORE $\underset{p\times 1}{a}$ IN MODO TALE CHE $VAR(y_1) = VAR(\tilde{X}a)$ OPPURE $VAR(Za)$ SIA MASSIMA

OBIETTIVO $\max_a VAR(\underset{n\times 1}{y_1}) = \max_a VAR(Za)$

CON IL VINCOLO $a^T a = 1$

OSSIA $\sum_{i=1}^{p} a_i^2 = 1$

$$VAR(y_1) = \frac{\sum_{i=1}^{n} (y_{i1} - \bar{y}_1)^2}{n-1}$$

$\bar{y}_1 = 0$

$$VAR(y_1) = \frac{\sum_{i=1}^{n} y_{i1}^2}{n-1} = \frac{y_1^T y_1}{n-1} = \frac{a^T Z^T Z a}{n-1}$$

$$\max_a VAR(y_1) = \max_a \frac{a^T Z^T Z a}{n-1} = \max_a \underset{1\times p}{a^T} \underset{p\times p}{R} \underset{p\times 1}{a}$$

DOBBIAMO CERCARE IL VETTORE $\underset{p\times 1}{a} = (a_1 \dots a_p)^T$ IN MODO TALE CHE LO SCALARE

$a^T R a$ SIA IL PIU' GRANDE POSSIBILE

DATO CHE $R = V \Lambda V^T$ CON $V^T V = I = V V^T$

$$a^T R a = a^T V \Lambda V^T a .$$

SE PONIAMO $b = V^T a$

SI NOTI CHE $b^T b = a^T V V^T a = a^T a = 1$

$$= b^T \Lambda b = \sum_{i=1}^{p} b_i^2 \lambda_i \leq \sum_{i=1}^{p} b_i^2 \lambda_{max}$$

$$= \lambda_{max} \sum_{i=1}^{p} b_i^2 = \lambda_{max} b^T b = \lambda_{max}$$

CONCLUSIONE: IL VALORE MASSIMO CHE PUO' ASSUMERE $VAR(y_1)$ E' λ_1 = IL PIU' GRANDE AUTOVALORE DELLA MATRICE $R(S)$

AFFINCHÉ $VAR(y_1) = \lambda_1$ COME DEVE ESSERE SCELTO a?

$$\underset{1\times p}{a^T} \underset{p\times p}{V} \Lambda V^T a = 1 \quad \text{SE}$$

$$\underset{1\times p}{a^T V} = (1, 0, \dots, 0) \qquad \underset{p\times 1}{V^T a} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\underset{1\times p}{(1, 0 \dots 0)} \underset{p\times p}{\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & & \lambda_p \end{pmatrix}} \underset{p\times 1}{\begin{pmatrix} 1 \\ 0 \\ \\ 0 \end{pmatrix}} = \lambda_1$$

$$V^T a = \begin{pmatrix} 1 \\ 0 \\ \\ 0 \end{pmatrix} \quad \text{SE} \quad a = V_1 = \text{AUTOVETTORE ASSOCIATO A } \lambda_1$$

$$V^T V_1 = \begin{pmatrix} V_1^T V_1 \\ V_2^T V_1 \\ \vdots \\ V_p^T V_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

DATO CHE $V^T V = I$

RIASSUMENDO: IL VETTORE a CHE CONTIENE I COEFFICIENTI DELLA COMBINAZIONE LINEARE DELLE VARIABILI ORIGINARIE CHE MASSIMIZZA LA VARIANZA È IL PRIMO AUTOVETTORE ASSOCIATO AL PRIMO AUTOVALORE DELLA MATRICE $R(S)$

IN GENERALE IL NOSTRO OBIETTIVO E'

TROVARE r VETTORI $\underset{p\times 1}{e_1}\ \underset{p\times 1}{e_2} \dots \underset{p\times 1}{e_r}$ (CIASCUNO

COMBINAZIONE LINEARE DELLE VARIABILI ORIGINARIE

IN MODO TA

$$y_1 = Ze_1 \qquad y_2 = Ze_2 \qquad y_r = Ze_r$$

$$\underset{n\times r}{Y} = \underset{n\times p}{Z}\,\underset{p\times r}{A} \qquad \underset{p\times r}{A} = (e_1, e_2, \dots, e_r)$$

IN MODO TALE CHE

$\operatorname{tr} \operatorname{cov}(Y)$ SIA MASSIMA

OBIETTIVO $\max_A \operatorname{tr} \operatorname{cov}(Y) = \max_A \operatorname{tr} \operatorname{cov}(ZA)$

CON IL VINCOLO CHE $\underset{r\times p}{A^T}\,\underset{p\times r}{A} = I_r$

$$\operatorname{cov}(ZA) = \frac{A^T Z^T Z A}{n-1} = A^T R A$$

$$\underset{r\times p}{A^T}\,\underset{p\times p}{R}\,\underset{p\times r}{A} = \begin{pmatrix} \underbrace{e_1^T R e_1}_{1\times 1} & \underbrace{e_1^T R e_2}_{1\times 1} & \dots & e_1^T R e_r \\ e_2^T R e_1 & e_2^T R e_2 & \dots & e_2 R e_r \\ \vdots & & & \\ e_r R e_1 & e_r R e_2 & \dots & e_r R e_r \end{pmatrix}$$

$r\times r$

$$\mathrm{tr}\,(\mathrm{cov}\,ZA) = \sum_{i=1}^{r} a_i^T R a_i$$

DATO CHE $a_i^T R a_i$ E' MASSIMA QUANDO

$$a_i = v_i \quad i = 1, 2, \dots, r$$

$$\mathrm{tr}\,\mathrm{cov}(Y) = \mathrm{tr}\,\mathrm{cov}(ZA)$$ E' MASSIMA QUANDO A CONTIENE I PRIMI r AUTOVETTORI DELLA MATRICE R

$$S = \begin{pmatrix} VAR(X_1) & cov(X_1 X_2) & & cov(X_1 X_p) \\ cov(X_1 X_2) & VAR(X_2) & \dots & cov(X_2 X_p) \\ \vdots & & \ddots & \vdots \\ cov(X_p X_1) & & & VAR(X_p) \end{pmatrix}$$

$$tr(S) = \sum_{i=1}^{p} VAR(X_i) = tr(V \Lambda V^T)$$
$$= tr(\Lambda V^T V) = tr(\Lambda) = \sum_{i=1}^{m} \lambda_i$$

CONCLUSIONE

1. LA SOMMA DEGLI AUTOVALORI DELLA MATRICE (S) E' UGUALE ALLA TRACCIA DELLA MATRICE S (SOMMA DELLE VARIANZE)

2. $$\frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{p} \lambda_i} = \frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{p} VAR(X_i)} =$$ QUOTA DI VARIANZA SPIEGATA DALLE PRIME r COMPONENTI PRINCIPALI

COROLLARIO: SE $S = R$
$$\sum_{i=1}^{p} VAR(X_i) = p$$

$$|S| = |V^T \Lambda V|$$
$$= |V^T| \, |\Lambda| \, |V|$$
$$= |\Lambda| = \prod_{i=1}^{p} \lambda_i$$

CONCLUSIONE IL DETERMINANTE DI UNA MATRICE E' UGUALE AL PRODOTTO DEI RISPETTIVI AUTOVALORI

(1)

(2) DATO CHE $|S|$ SI PUO' INTERPRETARE COME "VARIANZA GENERALIZZATA" L'ESTRAZIONE DELLE PRIME z PC MASSIMIZZA LA VARIANZA GENERALIZZATA CHE SI PUO' ESTRARRE CONSIDERANDO z COMBINAZIONI LINEARI DELLE VARIABILI ORIGINARIE

SCOMPOSIZIONE IN VALORI SINGOLARI

$\underset{n \times p}{X}$ DI RANGO $r \leq p$

$$\underset{n \times p}{X} = \underset{n \times r}{U} \, \underset{r \times r}{\Gamma} \, \underset{r \times p}{V^T}$$

$\underset{n \times r}{U}$ = MATRICE ORTONORMALE $= (u_1, \dots, u_r)$

$$\underset{r \times n}{U^T} \, \underset{n \times r}{U} = I_r$$

CHE CONTIENE GLI AUTOVETTORI DI XX^T

$\underset{p \times r}{V}$ = MATRICE ORTONORMALE $= (v_1, \dots, v_p)$

$$\underset{r \times p}{V^T} \, \underset{p \times r}{V} = I_r$$

CHE CONTIENE GLI AUTOVETTORI DI X^TX

$$\underset{r \times r}{\Gamma} = \begin{pmatrix} \gamma_1 & 0 & \dots & 0 \\ 0 & \gamma_2 & \dots & 0 \\ & & \ddots & \\ 0 & & \dots & \gamma_r \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \dots & 0 \\ & & \ddots & \\ 0 & \dots & & \sqrt{\lambda_r} \end{pmatrix}$$

Γ = MATRICE DIAGONALE CHE CONTIENE I VALORI SINGOLARI (RADICI QUADRATE DEGLI r AUTOVALORI NON NULLI DELLA MATRICE X^TX O XX^T)

$$\underset{n\times p}{X} = \sum_{i=1}^{r} u_i \gamma_i v_i^T$$

CONCLUSIONE $\underset{n\times p}{X}$ PUÒ ESSERE SCRITTA COME SOMMA DI r MATRICI DI DIMENSIONI $n\times p$ CIASCUNA DI RANGO 1

OBIETTIVO: AL POSTO DELLA MATRICE X VOGLIAMO SOSTITUIRE UNA MATRICE $\hat{X}$ IN MODO TALE CHE

$$\min_{\hat{x}_{ij}} \sum_{i=1}^{n} \sum_{j=1}^{p} \left(x_{ij} - \hat{x}_{ij}\right)$$

DATO CHE $\sum_{i=1}^{n} \sum_{j=1}^{p} x_{ij}^2 = tr(X'X)$ SI OTTIENE

$$\min_{\hat{x}_{ij}} tr\,(X-\hat{X})'(X-\hat{X})$$

SE LA MATRICE $\hat{X}$ HA RANGO 2

$$X-\hat{X} = \sum_{i=1}^{r} u_i \gamma_i v_i^T - \sum_{i=1}^{2} u_i \gamma_i v_i^T$$

$$= \sum_{i=3}^{r} u_i \gamma_i v_i^T$$

$$tr\,(X-\hat{X})'(X-\hat{X}) = tr\left(\sum_{i=3}^{r} u_i \gamma_i v_i^T\right)'\left(\sum_{j=3}^{r} u_j \gamma_j v_j^T\right)$$

$$= tr \sum_{i=3}^{r}\sum_{j=3}^{r} v_i \gamma_i u_i^T u_j \gamma_j v_j^T =$$

DATO CHE $u_i^T u_j = 0$ per $i \neq j$

$$= tr \sum_{i=3}^{r} v_i \gamma_i^2 v_i^T = tr \sum_{i=3}^{r} \gamma_i^2 v_i^T v_i$$

$$= tr \sum_{i=3}^{r} \gamma_i^2 = \sum_{i=3}^{r} \gamma_i^2 = \sum_{i=3}^{r} \lambda_i$$

QUINDI: LA MIGLIOR RAPPRESENTAZIONE DI RANGO 2 DELLA MATRICE ORIGINARIA (OSSIA QUELLA CHE MINIMIZZA LA SOMMA DEI QUADRATI DEI RESIDUI DELLE p VARIABILI) SI OTTIENE PRENDENDO I PRIMI DUE AUTOVETTORI ASSOCIATI AI PRIMI DUE AUTOVALORI

OSS: $\sum_{i=3}^{r} \lambda_i$ E' MINIMA QUANDO $\lambda_3, \dots, \lambda_r$ SONO I PIU' PICCOLI AUTOVALORI DELLA MATRICE $S(R)$