ANALISIS DE COMPONENTES PRINCIPALES

1. REPASO

- VALORES PROPIOS : λ_ℓ

$\underline{C}$ matriz de $L \times L$ elementos

$|\underline{C} - \lambda_\ell \underline{I}| = 0$ λ_ℓ: valores propios de $\underline{C}$
$\ell = 1, 2, \dots L$

Ejemplo: $\underline{C} = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$

$|\underline{C} - \lambda \underline{I}| = (4-\lambda)(2-\lambda) - 3 = 0$

$\lambda_1 = 5 \quad \lambda_2 = 1$

- VECTORES PROPIOS : $\underline{v}_\ell$

$\underline{C}\,\underline{v}_\ell = \lambda_\ell \underline{v}_\ell$ $\underline{v}_\ell$: vectores propios
$L \times 1$ elementos

Como hay infinitas soluciones : s.t. $\|\underline{v}_\ell\| = 1$

$\{\underline{v}_\ell\}$ son ortonormal : $\underline{v}_\ell^T \underline{v}_k = \begin{cases} 1 & \text{si } \ell = k \\ 0 & \ell \neq k \end{cases}$

Se define la matriz de vectores propios:

$$\underline{A} = [\, \underline{v}_1 \ \underline{v}_2 \ \dots \ \underline{v}_L \,] \quad \text{de } L \times L$$

$$\underline{A}\,\underline{A}^T = \underline{I}$$

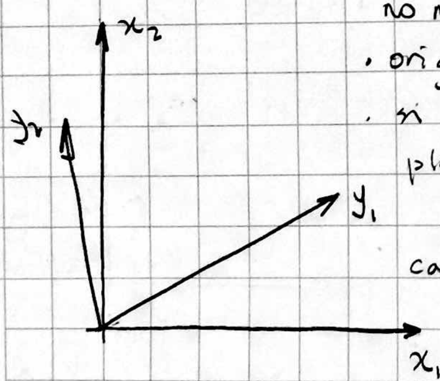
$$\underline{C}\,\underline{A} = [\, \lambda_1 \underline{v}_1 \ \ \lambda_2 \underline{v}_2 \ \dots \ \lambda_L \underline{v}_L \,]$$

⊙ TRANSFORMACIONES LINEALES

$$y_1 = a_{11} x_1 + a_{12} x_2 + \dots \ a_{1N} x_N$$
$$\vdots$$
$$y_N = a_{N1} x_1 + a_{N2} x_2 + \dots \ a_{NN} x_N$$

$$\underline{y} = \underline{A}\,\underline{x}$$

- ángulo entre y_1 e y_2 no necesariamente $\perp$
- origen coincide
- si se necesita desplazamiento agregar $+ a_{i0}$ en cada fila



⊙ PROMEDIO Y VARIANZA

$$\underline{x} = [\, x_1 \ x_2 \ \dots \ x_N \,]^T$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$\sigma_x = \frac{1}{N-1} \sum_{i=1}^{N} (\bar{x} - x_i)^2$$

$\frac{1}{N-1}$: unbiased estimator
→ ver Anexo

⊙ COVARIANZA

$$\underset{N \times L}{\underline{X}} = \begin{bmatrix} \underline{x}_1 & \underline{x}_2 & \dots & \underline{x}_L \end{bmatrix} = \begin{bmatrix} x_{11} & & \\ & \ddots & \\ & & x_{NL} \end{bmatrix}$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$\bar{\underline{X}} = [\, \bar{x}_1 \ \ \bar{x}_2 \ \ \dots \ \ \bar{x}_L \,]$$

$$\underline{X}_0 = \begin{bmatrix} x_{11} - \bar{x}_1 & & x_{1L} - \bar{x}_L \\ x_{21} - \bar{x}_1 & \dots & x_{2L} - \bar{x}_L \\ \vdots & & \\ x_{N-1} - \bar{x}_1 & & x_{NL} - \bar{x}_L \end{bmatrix}$$

$$= \underline{X} - \underline{M}_x \quad \text{donde} \quad M_x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \bar{\underline{X}}$$

La covarianza se define como:

$$\underline{C}_x = \frac{1}{N-1}\, \underline{X}_0^T \underline{X}_0$$

$$= \mathrm{cov}(\underline{X})$$

$$C_x(\ell,\ell) = \sigma^2_{x_\ell}$$

$C_x(\ell,k) =$ covarianza de $\underline{x}_\ell$ con $\underline{x}_k$

- > 0 dependencia lineal positiva
- $= 0$ no hay dependencia
- < 0 dependencia negativa

→ ver PAT03_CovarianceVisualization.m

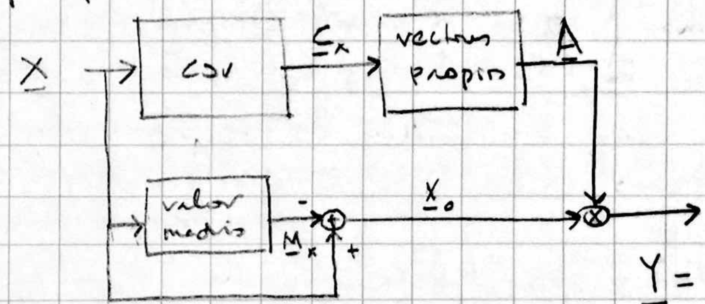
2. TRANSFORMACIÓN KARHUNEN-LOEVE

1947

$$\underline{X} = \left[\, \underline{x}_1 \;\; \underline{x}_2 \;\cdots\; \underline{x}_L \,\right] \qquad \underline{X}_0 = \underline{X} - \underline{M}_x$$

$$\underset{N\times L}{\underline{Y}} = \overset{N\times L}{\underline{X}_0}\;\overset{L\times L}{\underline{A}}$$

¿Qué pasa si $\underline{A}$ es la matriz de vectores propios de $\underline{C}_x$?



$$\underline{Y} = \underline{X}_0\, \underline{A}$$

$\underline{Y}$ es una transformación lineal de $\underline{X}_0$

Covarianza de $\underline{Y}$

$$\underline{C}_Y = \frac{1}{N-1}\, \underline{Y}_0^T \underline{Y}_0$$

$$\underline{Y}_0 = \underline{Y} - \underline{M}_Y$$

como $\underline{Y} = \underline{X}_0 \underline{A}$ y la media de las columnas de $\underline{X}_0$ es cero entonces $\underline{M}_Y = \underline{0}$

$$\underline{Y}_0 = \underline{Y} \Longleftarrow$$

entonces

$$\underline{C}_Y = \frac{1}{N-1} \underline{Y}^T \underline{Y}$$

$$= \frac{1}{N-1} (\underline{X}_o \underline{A})^T (\underline{X}_o \underline{A})$$

$$= \frac{1}{N-1} \underline{A}^T \underline{X}_o^T \underline{X}_o \underline{A}$$

como

$$\underline{C}_X = \frac{1}{N-1} \underline{X}_o^T \underline{X}_o$$

entonces

$$\underline{C}_Y = \underline{A}^T \underline{C}_X \underline{A}$$

como $\underline{A}$ son los valores propios de $\underline{C}_X$

$$\underline{C}_Y = [v_1 \; v_2 \cdots v_L]^T [\lambda_1 v_1 \cdots \lambda_L v_L]$$

$$= \begin{bmatrix} v_1^T \\ \vdots \\ v_L^T \end{bmatrix} [\lambda_1 v_1 \cdots \lambda_L v_L]$$

$$\underline{C}_Y = \begin{bmatrix} \lambda_1 v_1^T v_1 & \lambda_2 v_1^T v_2 & \cdots \\ & \ddots & \\ & & \lambda_L v_L^T v_L \end{bmatrix}$$

$$\underline{C}_Y = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \;\; \lambda_L \end{bmatrix}$$

Generalmente se ordena de tal forma que $\lambda_1 > \lambda_2 > \cdots > \lambda_L$

3. SIGNIFICADO

1. $\underline{Y}$ es una nueva representación de $\underline{X}$ (es una transformación LINEAL de $\underline{X}$).

2. Las columnas de Y no tienen dependencia lineal $C_Y(k,l) = 0$ si $k \neq l$.

3. Ordenado de mayor a menor, las columnas de Y están ordenadas por importancia.

→ ver PAT03_PCA_2D y PCA_SD.

4. APLICACIÓN A LA EXTRACCIÓN Y SELECCIÓN DE CARACTERÍSTICAS

$\underline{X}$ ← características extraídas NxL
(N muestras L características)

↓

$\boxed{\text{PCA}}$

↓

Y ← nuevas características NxL
las primeras p son las más importantes ← componentes principales

↓

Clasificador

Las deciciones se pueden tomar a partir de las primeras p características de $\underline{Y}$

ver → PAT03_PCA_FaceRecognition.m

Ventajas:
- La representación de muchas características es mucho más sencilla ya que se reduce la dimensionalidad (las gráficas son más simples).
- La clasificación usando las primeras componentes es más simple y más rápida para el clasificador.

Desventajas: Las características PCA dependen de todas las características extraídas (es necesario extraer todas las características para obtener Y, no hay ahorro de cómputo en la extracción de X)

ver → PAT03_PCA_DataCompression.m

BESSEL CORRECTION - ANEXO - | Fecha / Date:

Estimación de $\sigma^2 = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$

$$\bar{x} = \frac{1}{n}\sum x_i$$

$$E(s^2) = E\left\{\frac{1}{n-1}\sum(x_i-\bar{x})^2\right\}$$

$$= \frac{1}{n-1}E\left\{\sum(\overbrace{x_i-\mu}^{a}+\overbrace{\mu-\bar{x}}^{b})^2\right.$$

$$= \frac{1}{n-1}E\left\{\sum(x_i-\mu)^2+2(x_i-\mu)(\mu-\bar{x})+(\mu-\bar{x})^2\right\}$$

$$= \frac{1}{n-1}E\left\{\sum(x_i-\mu)^2+2(\mu-\bar{x})\sum(x_i-\mu)+n(\mu-\bar{x})^2\right\}$$

$$= \frac{1}{n-1}E\left\{\sum(x_i-\mu)^2-2(-\mu+\bar{x})(n\bar{x}-n\mu)+n(\mu-\bar{x})^2\right\}$$

$$= \frac{1}{n-1}E\left\{\sum(x_i-\mu)^2-2n(\bar{x}-\mu)^2+n(\bar{x}-\mu)^2\right\}$$

$$= \frac{1}{n-1}E\left\{\sum(x_i-\mu)^2-n(\bar{x}-\mu)^2\right\}$$

$$= \frac{1}{n-1}\left\{\sum E\{(x_i-\mu)^2\}-n\,E((\bar{x}-\mu)^2)\right\}$$

$$E\{(x_i-\mu)^2\} = \sigma^2$$

$$Var(\bar{x}) = E((\bar{x}-\mu)^2) = \frac{\sigma^2}{n}$$

$$E(s^2) = \frac{1}{n-1}\left(\sum\sigma^2-n\frac{\sigma^2}{n}\right) = \frac{1}{n-1}(n\sigma^2-\sigma^2) = \sigma^2$$

Notas / Notes: