Errores: ¿Cómo identificar los errores?

"$f \sim \chi^2$" $\leftarrow$ MSE es una función tipo χ^2

$$\chi^2_{OBS} - \chi^2_{OPT} \lesssim e_{TOLERANCIA}$$

$$\chi^2_{OBS} - \chi^2_{OPT} \simeq \sum_{jk} H_{jk} \Delta x_j \Delta x_k$$

con $\Delta x_j = x_{j\text{"OBS"}} - x_{j\text{"OPT"}}$

$$\Rightarrow (\Delta X)^2 = (\chi^2_{OBS} - \chi^2_{OPT}) \sum_{jk} (H^{-1})_{jk} \frac{\partial X}{\partial x_j} \frac{\partial X}{\partial x_k}$$

$$(\Delta X)^2 \lesssim e_{TOL} \underbrace{\sum_{jk} (H^{-1})_{jk}}_{\sum_{jk} (\Delta x_j \Delta x_k)} \frac{\partial X}{\partial x_j} \frac{\partial X}{\partial x_k}$$

$$\sum_{jk} (\Delta x_j \Delta x_k) \frac{\partial X}{\partial x_j} \frac{\partial X}{\partial x_k}$$

$$\boxed{\rightarrow \Delta x_j^2 = e_{TOL} (H^{-1})_{jj}}$$

Scipy se detiene cuando:

$$\frac{f_{k+1} - f_k}{\max(f_{k+1}, f_k, 1)} < e \quad \text{o tolerancia relativa.}$$

$$\Rightarrow \boxed{e_{TOL} = e \cdot \max(f_{k+1}, f_k, 1)}$$

Scanned with CamScanner

PCA = Análisis de Componentes Principales

Retomemos el problema original

$\mathbb{A}\mathbb{X} = \mathbb{B}$; $\mathbb{A}_{m \times n}$ con $m \gg n$ siendo m el # de observaciones y n el # de parámetros. ($\{p_1, p_2, \dots\}$)

$$\begin{array}{cccccc} & p_1 & p_2 & & p_n & \\ & \downarrow & \downarrow & & \downarrow & \\ \mathbb{A} = & \left(\begin{matrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \end{matrix} \right. & \begin{matrix} x_{12} \\ x_{22} \\ \vdots \\ x_{m2} \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \left. \begin{matrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{mn} \end{matrix} \right) & \\ & x_1 & x_2 & \cdots & x_n & \end{array}$$

$$\mathbb{X} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \quad \text{"pesos"}$$

$$\mathbb{B} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

La idea es describir linealmente

$$p_1 x_1 + p_2 x_2 + \dots + p_n x_n = b \quad \text{con } \{x_1, x_2, \dots, x_n ; b\} \text{ observaciones}$$

$\boxed{\Longrightarrow \text{encontrar los pesos.}}$

Sabemos que para la minimización, $\boxed{L(\mathbb{X}; \mathbb{A}, \mathbb{B})}$

$$\mathbb{X}_{OPT} = \underbrace{(\mathbb{A}^T \mathbb{A})^{-1}}_{\text{Examinemos esto}} \mathbb{A}^T \mathbb{B} \; ; \; L(\hat{\mathbb{B}}, \mathbb{B}) = \frac{1}{2m} \sum_{j=1}^{m} (b_j - \hat{b}_j)^2 \quad \text{con } \hat{b}_j = \sum_{k=1}^{n} x_{jk} p_k$$

¿Quién es $\frac{1}{m} \mathbb{A}^T \mathbb{A}$? Es la matriz de covarianza.

$$S = \frac{1}{m} \mathbb{A}^T \mathbb{A} \; ; \; S_{jk} = E(x_j x_k) = \sigma_{jk}^2 + \mu_j \mu_k$$

¿Cuál es el principio o la idea? Encontrar un conjunto de variables $\{y_i\}$ tal que maximicen la varianza.... ¿Porqué?

Si "normalizo" mi conjunto de muestras... i.e. $\tilde{x}_j = (x_j - \mu_j)/\sigma_j$

$$\tilde{A} = \begin{pmatrix} (x_{11}-\bar{x}_1)/\sigma_1 & (x_{12}-\bar{x}_2)/\sigma_2 & \cdots & (x_{1n}-\bar{x}_n)/\sigma_n \\ (x_{21}-\bar{x}_1)/\sigma_1 & (x_{22}-\bar{x}_2)/\sigma_2 & \cdots & (x_{2n}-\bar{x}_n)/\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ (x_{m1}-\bar{x}_1)/\sigma_1 & (x_{m2}-\bar{x}_2)/\sigma_2 & \cdots & (x_{mn}-\bar{x}_n)/\sigma_n \end{pmatrix}$$

$E(\{\tilde{x}_j\}) = 0$
$Vr(\{\tilde{x}_j\}) = 1$

Entonces

$$\underset{n\times n}{\tilde{S}} = \frac{1}{m}\tilde{A}^T A = \begin{pmatrix} \tilde{\sigma}_{11}^2 = 1 & \Delta_1\Delta_2 & \cdots & \Delta_1\Delta_n \\ \Delta_2\Delta_1 & \tilde{\sigma}_{22}^2 = 1 & \cdots & \Delta_2\Delta_n \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_n\Delta_1 & \Delta_n\Delta_2 & \cdots & \tilde{\sigma}_n^2 = 1 \end{pmatrix}$$

; La matriz de las desviaciones en cada variable.

Tiene varias propiedades:

1. Def. Positiva
2. Autovalores Todos Reales positivos ; $\lambda > 0$
3. Vectores propios son ortonormales
4. Si un Hessiano es PDM entonces la función es convexa
5. Simétrica
6. Det. es positivo
7. ¡Es una matriz de rotación!

⋮

Entre otras...

$$\tilde{S} = \hat{\mathbb{O}}\,\lambda\,\mathbb{O}^T$$

con $\hat{\mathbb{O}} = \begin{pmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \end{pmatrix}$ y $\lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & \lambda_n \end{pmatrix}$

$\boxed{q_j \text{ Rotaciones con Varianza } \lambda_j}$

¿Quién va a ser mi componente principal?

1. $Max\{\lambda_j\}$ y q_k asociado la combinación de "parámetros"
2. El siguiente max ... q_l asociado

⋮

r. Escojo el # que yo considere óptimo.