Capítulo 1

DVS. Geometría del ajuste.

1.1. Planteamiento general del ajuste en un espacio euclídeo ponderado.

El tratamiento del AFG, ACP y ACS del apartado anterior, supone obviar un planteamiento general en un espacio euclídeo ponderado, con métrica ponderada, puesto que aunque en ACS se han considerado las masas de los puntos, la forma cuadrática que define la métrica χ^2, como métrica euclídea ponderada, no ha sido considerada como tal, reconvirtiendo el problema a la métrica euclídea clásica, al no realizarse un tratamiento matricial. Para este planteamiento general vamos a considerar los siguientes elementos del problema:

- Sea una tabla Y de orden $I \times J$ y vamos a ajustar a la nube de I puntos fila, en el espacio J-dimensional, un subespacio óptimo.
- Sea $S \subset \mathbb{R}^J$ un subespacio K^*-dimensional de baja dimensión que se va a ajustar a la nube.
- Dado un punto cualquiera de la nube, y_i, con masa ω_i; sea $\widehat{y}_i$ el punto del subespacio S, que ajusta al y_i.
- El punto $\widehat{y}_i$ va a ser elegido de manera que la distancia d_i es:

$$d_i^2 = \| y_i - \widehat{y}_i \|_{\mathbb{D}_q}^2 = (y_i - \widehat{y}_i)' \mathbb{D}_q (y_i - \widehat{y}_i) \tag{1.1}$$

siendo $\mathbb{D}_q$ la matriz diagonal que define la métrica en el espacio $J-$dimensional.

- **La función criterio** para elegir el subespacio S es :

$$\Psi(S; y_1, \cdots, y_I) = \sum_i \omega_i d_i^2 = \sum_i \omega_i (y_i - \widehat{y}_i)' \mathbb{D}_q (y_i - \widehat{y}_i) \tag{1.2}$$

y **el criterio de optimación** es:

$$\min_S \Psi(S; y_1, \cdots, y_I) \tag{1.3}$$

Comentario 1.1.1 *El centroide de la nube en el J-espacio, $\bar{y}$, es el punto que minimiza*[1] *la función Ψ cuando el subespacio S es de dimensión $K^* = 0$, es decir:*

$$\min_s \left\{ \sum_i \omega_i (y_i - s)' \mathbb{D}_q (y_i - s) \right\}; \qquad s = \bar{y} \tag{1.4}$$

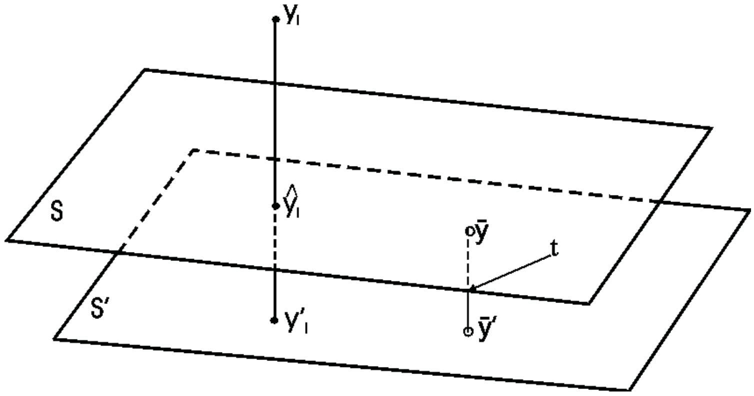
Demostración:

$$\frac{\partial \Psi}{\partial s} = -2 \sum_i \omega_i (y_i - s) \mathbb{D}_q = 0 \qquad \sum_i \omega_i y_i \mathbb{D}_q = \sum_i \omega_i s \mathbb{D}_q \qquad s = \frac{\sum_i \omega_i y_i}{\sum_i \omega_i} = \bar{y}$$

[1] La suma de los cuadrados de todos los valores respecto a su media es mínima.

Comentario 1.1.2 *El subespacio óptimo debe contener al centroide, por lo que se simplifica la búsqueda a aquellos que lo contienen.*

Figura 1.1: El subespacio óptimo debe contener al centroide



Demostración:
Sea $S^{'}$ un subespacio paralelo a S que no contiene al centroide $\bar{y}$, vamos a demostrar que no es óptimo en el sentido de mínimos cuadrados.
Sea $y_i^{'} \in S^{'}$ la proyección del punto y_i sobre $S^{'}$, se verifica:

$$\Psi(S^{'}; y_1, \ldots, y_I) = \sum_i \omega_i (y_i - y_i^{'})^{'} \mathbb{D}_q (y_i - y_i^{'})$$

Si suponemos que $\sum_i \omega_i = 1$, sea $\bar{y}^{'} \in S^{'}$ el punto más cercano a $\bar{y}$ y sea $t = \bar{y} - \bar{y}^{'}$ y $t = \hat{y}_i - y_i^{'}$, la distancia de la nube a $S^{'}$, la podemos expresar:

$$\Psi(S^{'}; y_1, \ldots, y_I) = \sum_i \omega_i (y_i - \hat{y}_i + \hat{y}_i - y_i^{'})^{'} \mathbb{D}_q (y_i - \hat{y}_i + \hat{y}_i - y_i^{'}) =$$

$$= \sum_i \omega_i (y_i - \hat{y}_i)^{'} \mathbb{D}_q (y_i - \hat{y}_i) + \sum_i \omega_i (\hat{y}_i - y_i^{'})^{'} \mathbb{D}_q (\hat{y}_i - y_i^{'}) + 2 \sum_i \omega_i (y_i - \hat{y}_i)^{'} \mathbb{D}_q (\hat{y}_i - y_i^{'}) =$$

$$= \sum_i \omega_i (y_i - \hat{y}_i)^{'} \mathbb{D}_q (y_i - \hat{y}_i) + t^{'} \mathbb{D}_q t$$

ya que el tercer término de la expresión anterior vale cero, por ser $\hat{y}_i - y_i^{'} = t$ un vector constante y cumplirse:

$$\sum_i \omega_i (y_i - \hat{y}_i) = \sum_i \omega_i y_i - \sum_i \omega_i \hat{y}_i = \bar{y} - \sum_i \omega_i (y_i^{'} + t) = \bar{y} - (\sum_i \omega_i y_i^{'} + t) = \bar{y} - (\bar{y}^{'} + t) = \bar{y} - \bar{y} = 0$$

Por lo tanto encontramos un subespacio S que contiene a $\bar{y}$ y que mejora la distancia al subespacio $S^{'}$ en $t^{'} \mathbb{D}_q t$.

Comentario 1.1.3 La formulación del problema *puede expresarse: Sea* $\{u_1, u_2, \cdots, u_{K^*}\}$ *una base del subespacio* S*, entonces se puede poner:* $\widehat{y}_i = \bar{y} + \sum_{k=1}^{K^*} f_{ik} u_k$ *y* **el criterio de optimización** *se puede formular:*

$$\min_S \Psi(S; y_1, \ldots, y_I) = \min_{\{u_1, \ldots, u_{K^*}\}} \left[\sum_i \omega_i \left(y_i - \bar{y} - \sum_{k=1}^{K^*} f_{ik} u_k \right)^{'} \mathbb{D}_q \left(y_i - \bar{y} - \sum_{k=1}^{K^*} f_{ik} u_k \right) \right] \tag{1.5}$$

por lo que la búsqueda de S se reduce a encontrar los K^* vectores $\{u_1, \cdots, u_{K^*}\}$ y a calcular $K^* \times J$ escalares, al estar en un espacio J-dimensional. Para identificar la base de vectores solución óptima del problema, de entre las infinitas bases K^* dimensionales aún restringiéndonos a las ortonormales, no vamos a recurrir a la diferenciación, sino a la Descomposición en Valores Singulares de una matriz.

1.2. D.V.S. Introducción.

La descomposición en valores singulares de una matriz es la factorización de una matriz en producto de otras tres, cuya estructura y geometría asociadas son fáciles de interpretar. Permiten en esencia, resolver los problemas de optimización al ajustar los subespacios factoriales a las nubes de puntos. El desarrollo de la DVS, lo haremos en espacios euclídeos y en espacios euclídeos ponderados.

1.3. DVS en el caso euclídeo.

Vamos a analizar la descomposición en valores singulares de una matriz (DVS), suponiendo de partida un espacio euclídeo, posteriormente extenderemos este concepto a un espacio euclídeo ponderado.

Definición 1.1 *Se llama DVS de la matriz real $A_{I\times J}$ a una factorización en tres matrices del tipo:*

$$A_{I\times J} = V_{I\times K}(\mathbb{D}_\alpha)_{K\times K}U'_{K\times J} \tag{1.6}$$

en donde

- $K = rg(A) \leq \text{mín}(I, J)$
- $\mathbb{D}_\alpha$ *es una matriz diagonal donde* $\alpha = (\alpha_1, \cdots, \alpha_K)'$ *son todos reales positivos. Estos valores se llaman valores singulares de la matriz* A.
- V *y* U *son ortonormales:* $V'V = U'U = I_K$. *A las columnas* v_k *de* $V_{I\times K}$, *se les denomina vectores singulares por la izquierda de* A, *y a las columnas* u_k *de* $U_{J\times K}$ *se les denomina vectores singulares por la derecha de* A.
- *Los vectores* v_k *forman una base ortonormal para las columnas de* A *en el espacio* $I-$*dimensional, y los* u_k *forman una base ortonormal para las filas de* A *en el espacio* $J-$*dimensional.*

Comentario 1.3.1 *Otra forma de escribir la DVS es*[2]*:*

$$A = \sum_k \alpha_k v_k u'_k \qquad k = 1, \cdots, K \qquad \alpha_1, \cdots, \alpha_k \geq 0 \tag{1.7}$$

es decir la DVS de A *se puede interpretar como una combinación lineal en términos de las matrices* $(v_k u'_k\ ;\ k = 1, \cdots, K)$ *que son de rango uno, donde los coeficientes* α_k, *ponderan la importancia del respectivo término* $k-$*ésimo matricial, en la representación de* A.

1.3.1. Autodescomposición de una matriz.

Definición 1.2 *La DVS para una matriz cuadrada real y simétrica*[3]*:* $B_{J\times J}$, *(donde* $rg(B) = K \leq J$*) es:*

$$B_{J\times J} = V_{J\times K}(\mathbb{D}_\lambda)_{K\times K}V'_{K\times J} = \sum_{k\in K} \lambda_k v_k v'_k \qquad \mathbb{D}_\lambda = diag(\lambda_1, \cdots ; \lambda_K)\ ;\ V'V = I \tag{1.8}$$

Los valores singulares son en este caso autovalores, y las matrices de vectores singulares a la izquierda y derecha coinciden y son los autovectores. Como B *es simétrica y real, sus autovalores son reales*[4].

[2]Esta expresión corresponde a la ya conocida de la reconstrucción de la matriz A donde $X = \sum_{\alpha=1}^{p} \sqrt{\lambda_\alpha} v_\alpha u'_\alpha$ siendo $\alpha_k = \sqrt{\lambda_k}$.

[3]B es ortogonalmente diagonalizable.

[4]La función criterio proviene de la autodescomposición de $X'X = U\mathbb{D}_\lambda U' \Longrightarrow U'X'XU = \mathbb{D}_\lambda$

1.4. Algunos resultados sobre la DVS.

Comentario 1.4.1 La existencia de la DVS de una matriz real $A_{I\times J}$ *se demuestra sin más que presuponer la existencia de las autodescomposiciones*[5] *de* $A^{'}A$ *y* $AA^{'}$*. Si se consideran dichas matrices tenemos:*

$$A = V\mathbb{D}_\alpha U^{'} \quad A^{'} = U\mathbb{D}_\alpha V^{'} \quad A^{'}A = U\mathbb{D}_\alpha V^{'}V\mathbb{D}_\alpha U^{'} = U\mathbb{D}_{\alpha^2}U^{'} \quad AA^{'} = V\mathbb{D}_\alpha U^{'}U\mathbb{D}_\alpha V^{'} = V\mathbb{D}_{\alpha^2}V^{'}$$

pero por la autodescomposición de $A^{'}A$ *y* $AA^{'}$ *tenemos*[6] *:*

$$A^{'}A = U\mathbb{D}_\lambda U^{'} \qquad AA^{'} = V\mathbb{D}_\lambda V^{'} \Longrightarrow \lambda_k = \alpha_k^2 \qquad \mathbb{D}_\alpha = \mathbb{D}_\lambda^{\frac{1}{2}} = \mathbb{D}_\mu^{\frac{1}{2}} \tag{1.9}$$

Comentario 1.4.2 *Si* U *está formada por columnas con los autovectores asociados a los autovalores* λ_k *de* $A^{'}A$ *y* V *está formada por columnas con los autovectores asociados a los* μ_k *autovalores de* $AA^{'}$*, entonces la DVS de* A *es:* $A = V\mathbb{D}_\alpha U^{'}$. **Conviene en este punto no confundir la DVS de** A **con las autodescomposiciones de** $A^{'}A$ **y** $AA^{'}$*, es decir podemos expresar la DVS de A:*

$$A_{I\times J} = V_{I\times K}(\mathbb{D}_\alpha)_{K\times K}U^{'}_{K\times J} \qquad V^{'}V = U^{'}U = I$$

en donde

- $\mathbb{D}_\alpha$*, valores singulares de A, son las raíces cuadradas de los autovalores de* $A^{'}A$ *o de* $AA^{'}$*.*
- *Las columnas de* V*, vectores singulares por la izquierda de A, son los autovectores de la matriz* $AA^{'}$*.*
- *Las columnas de* U*, vectores singulares por la derecha de A, son los autovectores de la matriz* $A^{'}A$*.*

Definición 1.3 Coordenadas de las filas y columnas de A*, respecto de los vectores base en* U *y* V*. Hemos dicho antes que los vectores singulares por la izquierda* v_k *(columnas de* V*) forman una base* $I-$*dimensional para las columnas de A, y que los* u_k*, columnas de* U*, forman una base* $J-$*dimensional para las filas de A, ambas ortonormalizadas, entonces:*

$$F_{I\times K} = V_{I\times K}(\mathbb{D}_\alpha)_{K\times K} \qquad G_{J\times K} = U_{J\times K}(\mathbb{D}_\alpha)_{K\times K} \tag{1.10}$$

estas matrices F *y* G *así definidas son precisamente las coordenadas de las filas y columnas de* A*.*

En efecto:

$$\begin{cases} A = V\mathbb{D}_\alpha U^{'} = FU^{'} \Longrightarrow F = AU \Longrightarrow a_i = \sum_{k=1}^K f_{ik}u_k \\ A' = U\mathbb{D}_\alpha V' = GV' \Longrightarrow G = A'V \Longrightarrow a_j = \sum_{k=1}^K g_{jk}v_k \end{cases} \tag{1.11}$$

siendo a_i la fila $i-$ésima de A (a_j la $j-$ésima columna de A). La fila $i-$ésima de F, (columna $j-$ésima de G), representa por tanto las coordenadas de a_i, (a_j) respecto de los vectores base u_k, (v_k). Recordemos aquí las expresiones de las coordenadas de las filas y columnas de una matriz X en AFG:

$$F = Xu_\alpha = \sqrt{\lambda_\alpha}v_\alpha \qquad G = X^{'}v_\alpha = \sqrt{\lambda_\alpha}u_\alpha$$

1.4.1. DVS completa.

Definición 1.4 *Interesa a veces para ciertos cálculos completar las bases ortonormales de* V *y* U *en sus respectivos espacios de forma que se obtengan matrices cuadradas:*

$$\widetilde{V} = (v_1, \cdots, v_K; \cdots, v_I) \ ; \ I\times I \qquad \widetilde{U} = (u_1, \cdots, u_K; \cdots, u_J) \ ; \ J\times J$$

donde $v_{K+1}, \cdots, v_I$ *son* $I-K$ *vectores ortonormales a* $v_1, v_2, \cdots, v_K$ *y* $u_{K+1}, \cdots, u_J$ *son* $J-K$ *vectores ortonormales a* $u_1, u_2, \cdots, u_K$ *de forma que :*

$$\widetilde{V}^{'}\widetilde{V} = I \ ; \ \widetilde{U}^{'}\widetilde{U} = I \ ; \ \mathbf{A_{I\times J}} = \widetilde{\mathbf{V}}_{\mathbf{I\times I}}\mathbf{\Delta_{I\times J}}\widetilde{\mathbf{U}}^{'}_{\mathbf{J\times J}} \ ; \ \textit{siendo } \Delta = \left(\begin{array}{c|c} \mathbb{D}_\alpha & 0 \\ \hline 0 & 0 \end{array}\right) \tag{1.12}$$

Hay infinitas formas de completar las bases.

[5] Las autodescomposiciones existen, al ser los productos $A^{'}A$ y $AA^{'}$ matrices reales y simétricas, pag 625 J. de Burgos.
[6] Este desarrollo corresponde con los resultados del método no matricial.

1.5. Aproximación de una matriz por otra de bajo rango mediante DVS.

Dada una matriz $A_{I\times J}$ real, la DVS se puede expresar:

$$A = \sum_{k}^{K} \alpha_k v_k u_k^{'} \; ; \; \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_K > 0 \; ; \; \text{(Fórmula de reconstrucción de la matriz)} \tag{1.13}$$

El concepto de aproximación de A surge, cuando a partir de un cierto K^* los valores singulares son pequeños comparados con los anteriores de forma que $\alpha_{K^*+1}, \alpha_{K^*+2}, \cdots, \alpha_K$ representan valores pequeños en relación a $\alpha_1, \alpha_2, \cdots, \alpha_{K^*}$. En este caso podemos considerar $A_{[K^*]} = \sum_{k=1}^{K^*} \alpha_k v_k u_k^{'}$ como una aproximación de menor rango de A.
En efecto si consideramos **la función objetivo**[7]:

$$\operatorname{tr}\left((A-X)(A-X)^{'}\right) = \sum_{i\in I}\sum_{j\in J}(a_{ij}-x_{ij})^2 \; ; \; A_{I\times J} \; ; \; X_{I\times J} \tag{1.14}$$

con **el criterio de minimizarla** (mínimos cuadrados) se verifica:

$$\min_{X\in\mathcal{F}} \operatorname{tr}\left((A-X)(A-X)^{'}\right) = \operatorname{tr}\left((A-A_{[K^*]})(A-A_{[K^*]})^{'}\right) \tag{1.15}$$

en donde $\mathcal{F}$ es el espacio de todas las matrices $(I\times J)$ de rango $\leq K^*$.

Comentario 1.5.1 La demostración *de este resultado es como sigue: Sea $A_{I\times J}$ y consideremos su DVS completa, entonces la función objetivo se puede poner*[8]*:*

$$\operatorname{tr}\left((A-X)(A-X)^{'}\right) = \operatorname{tr}\left(\widetilde{V}\widetilde{V}^{'}(A-X)\widetilde{U}\widetilde{U}^{'}(A-X)^{'}\right) = \operatorname{tr}\left[[\widetilde{V}\widetilde{V}^{'}A - \widetilde{V}\widetilde{V}^{'}X][\widetilde{U}\widetilde{U}^{'}A^{'} - \widetilde{U}\widetilde{U}^{'}X^{'}]\right] =$$

$$= \operatorname{tr}(\widetilde{V}\underbrace{\widetilde{V}^{'}A\widetilde{U}}_{\Delta}\widetilde{U}^{'}A^{'} - \widetilde{V}\underbrace{\widetilde{V}^{'}A\widetilde{U}}_{\Delta}\widetilde{U}^{'}X^{'} - \widetilde{V}\underbrace{\widetilde{V}^{'}X\widetilde{U}}_{\Theta}\widetilde{U}^{'}A^{'} + \widetilde{V}\underbrace{\widetilde{V}^{'}X\widetilde{U}}_{\Theta}\widetilde{U}^{'}X^{'}) = (^{9})$$

$$= \operatorname{tr}(\Delta\Delta^{'} - \Delta\Theta^{'} - \Theta\Delta^{'} + \Theta\Theta^{'}) = \operatorname{tr}\left((\Delta-\Theta)(\Delta-\Theta)^{'}\right) \qquad \textit{donde:}$$

$$\begin{cases} \Delta = \widetilde{V}^{'}A\widetilde{U} & \Delta^{'} = \widetilde{U}^{'}A^{'}\widetilde{V} \\ \Theta = \widetilde{V}^{'}X\widetilde{U} & \Theta^{'} = \widetilde{U}^{'}X^{'}\widetilde{V} \end{cases} \qquad \textit{y como} \qquad \begin{cases} A = \widetilde{V}\Delta\widetilde{U}^{'} & A^{'} = \widetilde{U}\Delta^{'}\widetilde{V}^{'} \\ X = \widetilde{V}\Theta\widetilde{U}^{'} & X^{'} = \widetilde{U}\Theta^{'}\widetilde{V}^{'} \end{cases} \qquad \Delta = \left(\begin{array}{c|c} \mathbb{D}_\alpha & 0 \\ \hline 0 & 0 \end{array}\right)$$

$$\textit{tenemos:} \qquad \operatorname{tr}\left((\Delta-\Theta)(\Delta-\Theta)^{'}\right) = \sum_{k=1}^{K}(\alpha_k - \theta_{kk})^2 + \sum_{k=1}^{K}\sum_{\substack{l=1\\ l\neq k}}^{K}\theta_{kl}^2 \; (^{10}) \tag{1.16}$$

---

[7] La traza del cuadrado de la diferencia entre dos matrices define el cuadrado de la distancia entre ambas, en sentido euclídeo.

[8] Por ser la DVS de A completa: $\widetilde{V}\widetilde{V}^{'} = I \qquad \widetilde{U}\widetilde{U}^{'} = I$.

[9] $\operatorname{tr}[\widetilde{V}(\Delta\widetilde{U}^{'}A^{'})] = \operatorname{tr}[(\Delta\widetilde{U}^{'}A^{'})\widetilde{V}] = \operatorname{tr}[\Delta(\widetilde{U}^{'}A^{'}\widetilde{V})] = \operatorname{tr}[\Delta\Delta^{'}]$.

[10]

$$\text{Si la matriz } \mathrm{A}_{3\times 4} \qquad \Delta - \Theta = \begin{pmatrix} \alpha_1 & 0 & 0 & 0 \\ 0 & \alpha_2 & 0 & 0 \\ 0 & 0 & \alpha_3 & 0 \end{pmatrix} - \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} \\ \theta_{31} & \theta_{32} & \theta_{33} & \theta_{34} \end{pmatrix}$$

$$\operatorname{tr}\left((\Delta-\Theta)(\Delta-\Theta)^{'}\right) = \operatorname{tr}\left[\begin{pmatrix} (\alpha_1-\theta_{11}) & -\theta_{12} & -\theta_{13} & -\theta_{14} \\ -\theta_{21} & (\alpha_2-\theta_{22}) & -\theta_{23} & -\theta_{24} \\ -\theta_{31} & -\theta_{32} & (\alpha_3-\theta_{33}) & -\theta_{34} \end{pmatrix}\begin{pmatrix} (\alpha_1-\theta_{11}) & -\theta_{12} & -\theta_{13} & -\theta_{14} \\ -\theta_{21} & (\alpha_2-\theta_{22}) & -\theta_{23} & -\theta_{24} \\ -\theta_{31} & -\theta_{32} & (\alpha_3-\theta_{33}) & -\theta_{34} \end{pmatrix}^{'}\right] =$$

$$= (\alpha_1-\theta_{11})^2 + \theta_{12}^2 + \theta_{13}^2 + \theta_{14}^2 + \theta_{21}^2 + (\alpha_2-\theta_{22})^2 + \theta_{23}^2 + \theta_{24}^2 + \theta_{31}^2 + \theta_{32}^2 + (\alpha_3-\theta_{33})^2 + \theta_{34}^2 =$$

$$= (\alpha_1-\theta_{11})^2 + (\alpha_2-\theta_{22})^2 + (\alpha_3-\theta_{33})^2 + \theta_{12}^2 + \theta_{13}^2 + \theta_{14}^2 + \theta_{21}^2 + \theta_{23}^2 + \theta_{24}^2 + \theta_{31}^2 + \theta_{32}^2 + \theta_{34}^2.$$

siendo rg$(\Theta) =$ *rg*(X)(¹¹) *y la matriz óptima* $X = \widetilde{V}\Theta\widetilde{U}'$*, al tener los mismos vectores singulares que* A*, tiene los mismos autovalores, siendo* Θ *una matriz diagonal de elementos* $\alpha_1, \cdots, \alpha_{K^*}$ *y los demás valores cero. Por tanto* $X = A_{[K^*]}$ *es la óptima, puesto que:*

$$\operatorname{tr}\Big((\Delta - \Theta)(\Delta - \Theta)^{'}\Big) = \sum_{k=K^*+1}^{K} \alpha_k^2$$

Comentario 1.5.2 Vamos a adaptar la notación *de la DVS de A cuando utilizamos una aproximación de bajo rango, en efecto:*

$$V = \left[V_{[K^*]} \mid V_{[K-K^*]}\right] \;;\; \mathbb{D}_\alpha = \left(\begin{array}{c|c} \mathbb{D}_{\alpha[K^*]} & 0 \\ \hline 0 & \mathbb{D}_{\alpha[K-K^*]} \end{array}\right) \;;\; U = \left[U_{[K^*]} \mid U_{[K-K^*]}\right]$$

$$A = A_{[K^*]} + \left[A - A_{[K^*]}\right] = V_{[K^*]}\mathbb{D}_{\alpha[K^*]}U^{'}_{[K^*]} + \underbrace{V_{[K-K^*]}\mathbb{D}_{\alpha[K-K^*]}U^{'}_{[K-K^*]}}_{\textit{matriz de residuos}} \tag{1.17}$$

es decir tenemos las matrices:

$$A = \sum_{k=1}^{K} \alpha_k v_k u_k^{'} \;;\; A_{[K^*]} = \sum_{k=1}^{K^*} \alpha_k v_k u_k^{'} \;;\; A - A_{[K^*]} = \sum_{k=K^*+1}^{K} \alpha_k v_k u_k^{'} \tag{1.18}$$

Luego las sumas de los elementos al cuadrado de estas matrices (es decir las trazas de cada matriz por su traspuesta), son:

$$\operatorname{tr}(AA^{'}) = \sum_{k=1}^{K} \alpha_k^2 = \operatorname{tr}(A^{'}A) \qquad \operatorname{tr}(A_{[K^*]}A^{'}_{[K^*]}) = \sum_{k=1}^{K^*} \alpha_k^2 \qquad \operatorname{tr}\Big((A - A_{[K^*]})(A - A_{[K^*]})^{'}\Big) = \sum_{k=K^*+1}^{K} \alpha_k^2 \tag{1.19}$$

Se suele tomar en ACP y en Análisis Datos como medida de la calidad de la aproximación por bajo rango, la tasa de inercia:

$$\tau_{K^*} = \frac{\sum_{k=1}^{K^*} \alpha_k^2}{\sum_{k=1}^{K} \alpha_k^2} \tag{1.20}$$

1.6. DVS en el caso euclídeo ponderado.

Hasta ahora hemos supuesto espacios euclídeos con métrica euclídea clásica, vamos ahora a considerar espacios euclídeos ponderados. Supongamos que en los espacios de dimensiones I, J de la matriz $A_{I\times J}$, tenemos dos matrices cuadradas: $\Omega_{I\times I}$ y $\Phi_{J\times J}$, definidas positivas y simétricas, tales que la normalidad de un vector n , $I-$dimensional por ejemplo, significa que $n^{'}\Omega n = 1$.

Definición 1.5 DVS generalizada: *Se comprueba fácilmente que existe una DVS extendida a esta situación métrica más general cuya expresión es:*

$$A_{I\times J} = N_{I\times K}(\mathbb{D}_\alpha)_{K\times K}M^{'}_{K\times J} = \sum_{k=1}^{K} \alpha_k n_k m_k^{'} \qquad r(A) = K \tag{1.21}$$

donde:

- $\mathbb{D}_\alpha$ *es una matriz diagonal de valores:* $\alpha_1 \geq \alpha_2 \geq, \cdots, \alpha_k > 0$ *llamados valores singulares generalizados de* A*.*
- n_k*, columnas de* N*, cumplen* $N^{'}\Omega N = I$ *y son los vectores singulares generalizados por la izquierda de* A*.*

[11] $\Theta = \widetilde{V}^{'}X\widetilde{U}$ y $X = \widetilde{V}\Theta\widetilde{U}^{'}$ (DVS de X), luego Θ debe ser diagonal y $\operatorname{rg}(\Theta) = \operatorname{rg}(X)$.

- *m_k, columnas de M, cumplen $M'\Phi M = I$ y son los vectores singulares generalizados por la derecha de A.*
- *$\Omega_{I\times I}$ y $\Phi_{J\times J}$ son cuadradas, definidas positivas y simétricas.*

Comentario 1.6.1 Demostración: *Sea la DVS de la matriz: $(\Omega^{\frac{1}{2}}A\Phi^{\frac{1}{2}})_{I\times J}$, es decir:*

$$\Omega^{\frac{1}{2}}A\Phi^{\frac{1}{2}} = V\mathbb{D}_\alpha U' \qquad V'V = U'U = I \qquad A = \Omega^{-\frac{1}{2}}V\mathbb{D}_\alpha U'\Phi^{-\frac{1}{2}} \tag{1.22}$$

definimos:

$$N = \Omega^{-\frac{1}{2}}V \; ; \; M = \Phi^{-\frac{1}{2}}U \Longrightarrow N\mathbb{D}_\alpha M' = \Omega^{-\frac{1}{2}}V\mathbb{D}_\alpha U'\Phi^{-\frac{1}{2}} = \Omega^{-\frac{1}{2}}\Omega^{\frac{1}{2}}A\Phi^{\frac{1}{2}}\Phi^{-\frac{1}{2}} = A$$

y además se verifica:

$$N'\Omega N = V'\Omega^{-\frac{1}{2}}\Omega\Omega^{-\frac{1}{2}}V = V'V = I \; ; \; M'\Phi M = U'\Phi^{-\frac{1}{2}}\Phi\Phi^{-\frac{1}{2}}U = U'U = I$$

Definición 1.6 Coordenadas de las filas y columnas de A*, respecto de los vectores base en M y N.*

$$F_{I\times K} = N_{I\times K}(\mathbb{D}_\alpha)_{K\times K} = \Omega^{-\frac{1}{2}}V\mathbb{D}_\alpha \qquad G_{J\times K} = M_{J\times K}(\mathbb{D}_\alpha)_{K\times K} = \Phi^{-\frac{1}{2}}U\mathbb{D}_\alpha \tag{1.23}$$

estas matrices F y G son las coordenadas de las filas y columnas de A respecto a los vectores de las bases en M y N ya que:

$$\begin{cases} A = N\mathbb{D}_\alpha M' = FM' \Longrightarrow F = A\Phi M \Longrightarrow a_i = \sum_{k=1}^{K} f_{ik}m_k \\ A' = M\mathbb{D}_\alpha N' = GN' \Longrightarrow G = A'\Omega N \Longrightarrow a_j = \sum_{k=1}^{K} g_{jk}n_k \end{cases} \tag{1.24}$$

Definición 1.7 Aproximación de bajo rango en el caso generalizado. *Si eliminamos los últimos A_{K-K^*} términos de $A = \sum_k^K \alpha_k n_k m_k'$, la expresión siguiente:*

$$A_{[K^*]} = \sum_{k=1}^{K^*}\alpha_k n_k m_k' = N_{[K^*]}\mathbb{D}_\alpha M_{[K^*]}' \tag{1.25}$$

es la aproximación generalizada de A de rango K^*, que es mínimo cuadrática, es decir minimiza:

$$\min_{X\in\mathcal{F}} \operatorname{tr}\left(\Omega(A-X)\Phi(A-X)'\right) \tag{1.26}$$

donde $\mathcal{F}$ es el espacio de todas las matrices $I\times J$ de rango $\leq K^*$.

En la práctica se elige la dimensión K^* del subespacio en función de que presente una clara diferencia entre α_{K^*} y α_{K^*+1}.

1.7. La DVS y la solución del problema de ajuste de un subespacio.

Habíamos visto en el problema de optimización planteado antes, que el subespacio que se trata de ajustar a la nube de puntos dada $(y_i; i = 1,\cdots,I)$, se obtiene determinando un conjunto de ejes $\{u_1,\cdots,u_{K^*}\}$ en el espacio $J-$dimensional de forma que determinen un subespacio S óptimo, en el sentido de minimizar la expresión (1.5):

$$\min_{\{u_1,\cdots,u_{K^*}\}}\left[\sum_i \omega_i\left(y_i - \bar{y} - \sum_{k=1}^{K^*} f_{ik}u_k\right)'\mathbb{D}_q\left(y_i - \bar{y} - \sum_{k=1}^{K^*} f_{ik}u_k\right)\right]$$

estas expresiones dependen de:

1. Las masas w_i asignadas a cada punto y_i ; $i = 1,\cdots,I$.

2. La métrica euclídea ponderada asociada al $J-$espacio, definida por $(\mathbb{D}_q)_{J\times J}$ de pesos, cuya diagonal contiene las ponderaciones de cada dimensión del espacio $J-$dimensional.

Comentario 1.7.1 *La descomposición:*

$$A_{I\times J} = N_{I\times K}(\mathbb{D}_\alpha)_{K\times K}M'_{K\times J} = \sum_{k=1}^{K}\alpha_k n_k m'_k \qquad N'\Omega N = M'\Phi M = I$$

no es más que la DVS **no** generalizada, adaptada a una estructura de espacio euclídeo ponderado*, ya que las matrices Ω y Φ respecto de las que ortonormalizamos, no son más que las que definen las métricas de masas y pesos, en los espacios I y J dimensionales.*

Comentario 1.7.2 *Se puede considerar que la búsqueda de un subespacio óptimo S determinado por los ejes $\{u_1, \cdots, u_{K^*}\}$ es un caso particular de la optimización conseguida por la DVS generalizada.*

En efecto, cuando Ω es la matriz diagonal $\mathbb{D}_w$ de las masas de los puntos y cuando Φ es la matriz diagonal $\mathbb{D}_q$ de las ponderaciones asociadas a cada dimensión del espacio $J-$dimensional, la expresión a minimizar se escribe:

$$\text{tr}\left(\mathbb{D}_w(A-X)\mathbb{D}_q(A-X)'\right) = \sum_{i=1}^{I} w_i(a_i - x_i)'\mathbb{D}_q(a_i - x_i) \tag{1.27}$$

donde:

1. $(\mathbb{D}_w)_{I\times I} = \text{diag}(w_1, \cdots, w_I)$; $w_i =$ masas $i = 1, \cdots, I$
2. $(\mathbb{D}_q)_{J\times J} = \text{diag}(q_1, \cdots, q_J)$; $q_j =$ pesos $j = 1, \cdots J$
3. $a_i =$ fila $i-$ésima de A escrita como columna, $(a_i)_{J\times 1}$; $i = 1, \cdots, I$
4. $x_i =$ fila $i-$ésima de X escrita como columna, $(x_i)_{J\times 1}$; $i = 1, \cdots, I$

Por lo tanto podemos resolver el problema de optimización, mediante la DVS generalizada, de forma que la solución es:

$$X = A_{[K^*]} = \sum_{k=1}^{K^*}\alpha_k n_k m'_k = N_{[K^*]}\mathbb{D}_{\alpha[K^*]}M'_{[K^*]}$$

verificándose que:

1. Los vectores $m_1, \cdots, m_{[K^*]}$ definen los ejes factoriales o ejes principales, es decir el subespacio ajustado $S \subset \mathbb{R}^J$ que queríamos obtener, para las filas de la matriz A.
2. Los vectores $n_1, \cdots, n_{[K^*]}$ definen los ejes factoriales del subespacio ajustado K^*-dimensional $S' \subset \mathbb{R}^I$ para las columnas de la matriz A, mediante un procedimiento similar al utilizado para las filas de la matriz.
3. Las filas de $F_{[K^*]} = \left(N_{[K^*]}\mathbb{D}_{\alpha[K^*]}\right)_{I\times K^*} = \left[\alpha_1 n_1 \mid \alpha_2 n_2 \mid, \cdots, \mid \alpha_{[K^*]}n_{[K^*]}\right]$ son las coordenadas de los puntos-filas proyectados de la nube, sobre la estructura factorial anterior $m_{[K^*]}$.
4. Las filas de $G_{[K^*]} = \left(M_{[K^*]}\mathbb{D}_{\alpha[K^*]}\right)_{J\times K^*} = \left[\alpha_1 m_1 \mid \alpha_2 m_2 \mid, \cdots, \mid \alpha_{[K^*]}m_{[K^*]}\right]$ son las coordenadas de los puntos-columnas proyectados de la nube, sobre la estructura factorial anterior $n_{[K^*]}$.
5. α_k ; $k = 1, \cdots, K^*$, son los K^* primeros valores singulares de A en orden decreciente.

Comentario 1.7.3 *Comparando (1.5) con (1.27) observamos que la matriz A, se define como la matriz de filas centradas de Y: $A = Y - 1\bar{y}'$:*

$$\sum_{i=1}^{I} w_i(a_i - x_i)'\mathbb{D}_q(a_i - x_i) = \left[\sum_i \omega_i\left(\underbrace{y_i - \bar{y}}_{a_i} - \sum_{k=1}^{K^*} f_{ik}u_k\right)'\mathbb{D}_q\left(y_i - \bar{y} - \sum_{k=1}^{K^*} f_{ik}u_k\right)\right]$$

donde $\bar{y}$ es el centroide de la nube de puntos fila de A.

Por lo que la matriz A de la DVS generalizada (1.21), es una matriz centrada.

Definición 1.8 Representación Biplot. *Las filas de* $F_{[K^*]} = N_{[K^*]}\mathbb{D}_{\alpha[K^*]}$ *se representan en el subespacio* S*, normalmente de 2 o 3 dimensiones, para observar la configuración multidimensional de las filas de la matriz de datos* A*, y las filas de* $G_{[K^*]} = M_{[K^*]}\mathbb{D}_{\alpha[K^*]}$ *se representan en el subespacio* S'*, para observar la configuración multidimensional de las columnas de* A*. Sin embargo es posible representar las filas de* $G_{[K^*]}$ *como* $G_{[K^*]} = M_{[K^*]}$ *en el mismo espacio que el de las filas de* $F_{[K^*]}$*, esta particular representación común de puntos fila y columna de la matriz* A*, se llama representación biplot.*

En una representación biplot $A_{[K^*]} = F_{[K^*]}G'_{[K^*]}$ luego el producto escalar de la i-ésima fila de F y la j-ésima fila de G, es un valor aproximado del elemento a_{ij} de la matriz:

$$a_{ij} \approx f_i' g_j = |f_i| \cdot |g_j| \cos(f_i, g_j) \tag{1.28}$$

Los datos a_{ij} se centran usualmente, por ejemplo como en 1.7.3 respecto a la media de las columnas, y así una desviación $a_{ij} > 0$ se indica por vectores f_i y g_j formando un ángulo agudo y $a_{ij} < 0$ por vectores bajo un ángulo obtuso.