

## MODÈLE DE DÉCOMPOSITION MULTIDIMENSIONNELLE (G11, K1, K11)

(10 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

Le **modèle de décomposition multidimensionnelle** est un **modèle statistique** utilisé notamment en psychologie (psychométrie). Il est alors appelé **modèle de différenciations individuelles**.

(i) Soit  $T = (t_l)_{l \in \mathcal{J}}$  un **tableau statistique multidimensionnel**, avec  $t_l \in \mathbf{R}$ ,  $\forall l = (i_1, \dots, i_k) \in \mathcal{J}$ , où  $\mathcal{J} = \prod_{\alpha=1}^k \mathcal{J}_\alpha$  et  $\mathcal{J}_\alpha = \{1, \dots, n_\alpha\}$ ,  $\forall \alpha \in N_k^*$ .

On appelle **modèle de décomposition multidimensionnelle**, ou encore **modèle à multi-modalités**, ou encore **modèle à plusieurs modalités**, un modèle à  $k$  dimensions qui consiste à décomposer  $T$  selon la forme générale suivante, dite **forme de L.R. TUCKER** :

$$(1) \quad t_l = \sum_{j \in \mathcal{J}} \left\{ \prod_{\alpha=1}^k a_{\alpha, i(\alpha), j(\alpha)} \right\} \cdot c_j + u_l, \quad \forall j = (j_1, \dots, j_k) \in \mathcal{J},$$

où l'on note, par commodité,  $i(\alpha)$ ,  $j(\alpha)$ , etc, au lieu de  $i_\alpha$ ,  $j_\alpha$ , etc.

Dans cette décomposition :

(a) les termes  $c_j$  définissent un autre **tableau multidimensionnel**  $C = (c_l)_{l \in \mathcal{J}}$ , de même nature que  $T$ , appelé **noyau de la décomposition** (1) ;

(b) les  $k$  **matrices**  $A_\alpha$  définies par :

$$(2) \quad A_\alpha = (a_{\alpha, i(\alpha), j(\alpha)})_{(i(\alpha), j(\alpha)) \in \mathcal{J}_\alpha^2}, \quad \forall (i_\alpha, j_\alpha) \in \mathcal{J}_\alpha^2,$$

sont appelées **composantes (matricielles)** de  $T$  dans la décomposition (1) ;

(c) le tableau (**inobservable**)  $U = (u_l)_{l \in \mathcal{J}}$  est assimilé à une **perturbation aléatoire**.

(ii) En général, la décomposition (1) n'est pas unique. Pour qu'elle le soit (**identifiabilité**), on impose notamment des conditions d'orthonormalité entre les composantes de  $T$  (cf **orthogonalité**), ie :

$$(3) \quad A_\alpha' A_\alpha = I_n, \quad \forall \alpha \in N_k^* \quad (\text{matrices unités})$$

(eg  $A_\alpha = I_{i(\alpha)}$  pour certains termes d'indices  $\alpha$ ).

(iii) En pratique, l'interprétation et l'**estimation** de (1) nécessitent généralement de simplifier cette décomposition.

En particulier, lorsque  $k = 2$ , ce type de modèle est voisin, dans son principe, de certains modèles d'**analyse factorielle** (eg **analyse en facteurs communs et spécifiques**).

L'estimation de (1) par la **méthode des moindres carrés ordinaires** (avec contraintes sur les  $A_\alpha$ ) s'effectue en minimisant la fonction matricielle (quadratique) :

$$(4) \quad q(A_1, \dots, A_k, C) = \|U\|^2, \quad \text{avec } \|U\|^2 = \sum_{l \in \mathcal{I}} u_l^2.$$

(iv) Le modèle précédent intervient notamment dans l'analyse des **échelles multidimensionnelles**.

Ainsi, lorsque  $k = 3$ ,  $\mathcal{I}_1$  représente un ensemble d'**unités statistiques** (individus),  $\mathcal{I}_2$  une **échelle** de valeurs (ie une **échelle de mesure**, cardinale ou ordinale) et  $\mathcal{I}_3$  un ensemble de **caractères** (ou **attributs**) ;  $t_l$  est alors eg un **score** quantitatif (eg « résultat » ou « note »). Dans ce cas, on a :

$$(5) \quad T = A_1 C (A_2' \otimes A_3'),$$

où  $T$  et  $C \in M_{n(1), n(2)n(3)}(\mathbf{R})$  sont les  $(n_1, n_2 \cdot n_3)$ -tableaux  $T$  et  $C$ , réarrangés selon l'**ordre** défini par le produit cartésien  $\mathcal{I}_1 \times (\mathcal{I}_2 \times \mathcal{I}_3)$ , et  $A \otimes B$  désigne le **produit tensoriel** de  $A$  et  $B$ .