SCHÉMA DE DUALITÉ (K4, K5)

(09 / 01 / 2020, © Monfort, Dicostat2005, 2005-2020)

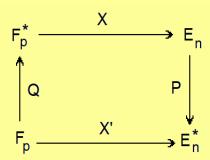
Le **schéma de dualité** est un **schéma** de base important de l'analyse linéaire des données.

(i) Soit $X = (x_{ij})_{(i,j)} \in M_{np}$ (R) une **matrice** représentant les **mesures** x_{ij} de p caractères statistiques (ou variable aléatoire) réel(le)s $j \in \{1,...,p\}$ = J effectuées sur des unités statistiques (ou individus) $i \in \{1,...,n\}$ = I.

L'espace des variables, ou espace des caractères, est $E_n = Im \ X$ et l'espace des unités statistiques est $F_p = Im \ X'$, avec $E_n \triangleleft R^n$ et $F_p \triangleleft R^p$ (sous espaces vectoriels). En effet :

- (a) \forall i \in I, X_i = g (i) \in \mathbb{R}^p est la i-ième ligne de X;
- (b) $\forall j \in J$, $x_j = h(j) \in \mathbb{R}^n$ est la j-ième colonne de X.

Soit E_n^* l'espace **dual** (algébrique) de E_n et F_p^* celui de F_p . Si E_n (resp F_p) est muni d'une **métrique** euclidienne de **matrice** P (resp Q) (cf **espace euclidien**), le **schéma de dualité** (cf ci-après) exprime alors une propriété de base de l'algèbre linéaire.



(ii) Ce schéma s'exprime, de façon équivalente, selon :

$$(1) X' = P X Q.$$

en identifiant les matrices et les **applications linéaires** auxquelles elles sont associées dans des **bases** données.

L'interprétation usuelle de la matrice P est celle d'une métrique euclidienne définie à partir d'une distance d_P entre caractères (cf aussi distance entre variables aléatoires):

(2)
$$d_{P}^{2}(x_{i'}, x_{j''}) = (x_{i'}, x_{j''})' P(x_{i'}, x_{j''}), \quad \forall (j', j'') \in J^{2}.$$

De même, la matrice Q est celle d'une métrique euclidienne définie à partir d'une distance d_Q entre individus) :

1

(3)
$$d_Q^2(x_{i'}, x_{i''}) = (x_{i'}, x_{i'})' P(x_{i'}, x_{i''}), \forall (i', i'') \in I^2.$$

On peut aussi interpréter P (resp Q) comme un isomorphisme entre E_n et E_n^* (resp entre F_p et F_p^*) (cf **homomorphisme**).