# MATRICES

## I. Bases de l'algèbre linéaire et modélisation en santé :

**Qu'est ce que modéliser ? Pourquoi le faire en santé ?**

- Modéliser permet de reproduire informatiquement une situation dans le but de tester des scénarios, des configurations…

**Qu'est ce qu'une matrice ?**

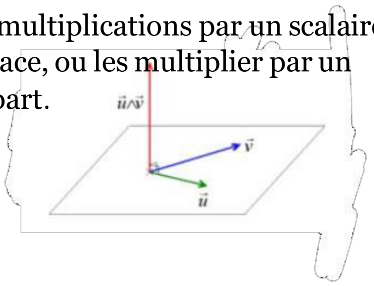
- Les matrices sont des grands tableaux avec lesquels ont peut effectuer des opérations mathématiques.
- Le calcul matriciel est précieux dans le domaine des statistiques multivariées.
  - Exemple : estimer l'effet d'une augmentation du prix des cigarettes sur le système de santé.

**Qu'est ce que l'algèbre linéaire ?**

- C'est le domaine des mathématiques qui étudie les transformations linéaires et les espaces vectoriels.
  - Ce sont des outils pour l'analyse multivariée de données dans un contexte de Big Data.

**Qu'est ce qu'un espace vectoriel ?**

- C'est une structure stable par addition de vecteurs et par multiplications par un scalaire.
- Autrement dit, on peut ajouter deux éléments d'un tel espace, ou les multiplier par un nombre, le résultat appartiendra encore à l'espace de départ.




## II. Le calcul matriciel

Une matrice est un tableau de nombres à n lignes et p colonnes : A (n, p)

C'est par exemple la donnée de n individus mesurés selon p variables (p≥1).

- Si p=1, on parle de univarié (= matrice colonne)

  Exemple : $A = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$

- si p ≥2, on parle de multivarié

  Exemple : $B = \begin{pmatrix} 5 & 1 & 6 \\ 3 & 4 & 2 \end{pmatrix}$

## Matrice identité :

- C'est une matrice **carrée** avec **tous les coefficients diagonaux égaux à 1** et tous les autres coefficients **nuls**.

Exemple : $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est la **matrice identité d'ordre 3**

## Matrices carrées :

- Si n=p, on parle de matrice carrée d'ordre n. Elle a autant de lignes (n) que de colonnes (p=n)

Exemple : $C = \begin{pmatrix} 3 & 3 \\ 9 & 2 \end{pmatrix}$ est une **matrice carrée d'ordre 2**

## Transposée de matrice :

- La transposée d'une matrice revient à présenter l'information qui est en colonnes en lignes et inversement.
- Ainsi la transposée de A(n,p), notée ${}^tA$ est une matrice à p lignes et n colonnes.
  Autrement dit, on s'intéresse à p variables déclinées selon n individus.

Le produit ${}^tA*A$ donne une matrice carrée d'ordre p, qui est une matrice symétrique (par rapport à la diagonale qui part du coefficient en haut à gauche de la matrice). +++

Exemple : $A = \begin{pmatrix} 2 & 1 & 5 \\ 6 & 1 & 3 \end{pmatrix}$ ${}^tA = \begin{pmatrix} 2 & 6 \\ 1 & 1 \\ 5 & 3 \end{pmatrix}$

**Point tut'** : La transposée d'une matrice existe toujours.

- Soit A une matrice carrée d'ordre n.
  1) A est dite **symétrique** si et seulement si ${}^tA = A$
  2) A est dite **antisymétrique** si et seulement si ${}^tA = -A$

Exemple : $\begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ est Symétrique

$\begin{pmatrix} 0 & -5 \\ 5 & 0 \end{pmatrix}$ est antisymétrique

## Matrice nilpotente :

- Une matrice est dite nilpotente d'ordre n lorsque $A^n = 0$ et $A^{n-1} \neq 0$

Une matrice est nilpotente lorsqu'il existe une puissance pour laquelle cette matrice est égale à la matrice nulle.

Exemple : $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

## Le déterminant d'une matrice :

- Pour une matrice d'ordre 2 :

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

Exemple : $\det \begin{pmatrix} 2 & 3 \\ 1 & 6 \end{pmatrix} = 12 - 3 = 9$

- Pour une matrice d'ordre 3 :

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \times \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \times \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \times \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

Exemple : $\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 \times \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} - 2 \times \det \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix} + 3 \times \det \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} = 1 \times (-3) - 2 \times (-6) + 3 \times (-3) = -3 + 12 - 9 = 0$

- Le déterminant sera utile pour calculer l'inverse

## Inverse de matrice :

- Est définie comme $A^{-1}$ tel que **$A.A^{-1} = I$** ou I est la **matrice identité**
- Elle n'existe que pour des matrices **carrées** si **$\det A \neq 0$**

**Point tut'** : La transposée d'une matrice inversible est toujours inversible

Soit A = $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice carrée d'ordre 2 :
Si det A ≠ 0 , alors :

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

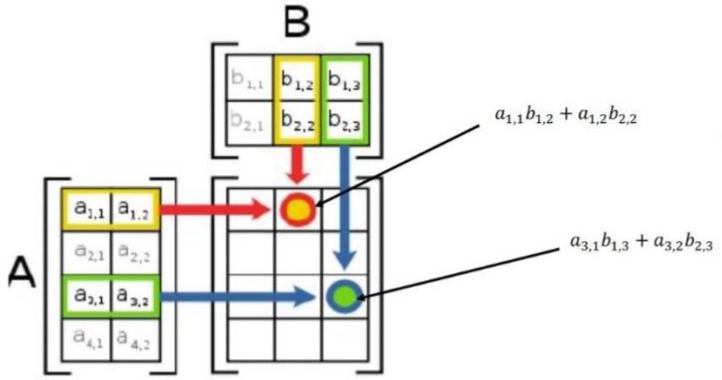
## Produits de matrices :

Pour calculer le produit de deux matrices A et B, il faut que le nombre de lignes de la deuxième matrice soit égale au nombre de colonnes de la première matrice.

Ainsi $A(n,p)*B(p, m) = C(n, m)$

On peut avoir un produit de matrice nul sans que l'une des matrices soit nulle

Méthode de calcul de A*B :




**Point Tut'** : Il n'est pas toujours facile de faire le produit de deux matrices, pour s'aider on peut placer les matrices de la façon suivante :

$$\begin{array}{cc} & \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ \begin{pmatrix} 5 \\ 6 \end{pmatrix} & \blacksquare \end{array}$$

Le but étant d'avoir un carré en bas à droite en plaçant la première matrice a gauche et la deuxième en haut comme dans le cas ci dessus. Si on n'obtient pas de carré alors on sait qu'on ne peut pas calculer le produit

Exemple : soit A = $\begin{pmatrix} 1 & 5 & -2 \\ 3 & 1 & 1 \end{pmatrix}$ et B = $\begin{pmatrix} -1 & 1 & 3 \\ 8 & 5 & 2 \\ 2 & 2 & 7 \end{pmatrix}$ On calcula A*B :

**AB = C =**

$$\begin{array}{cc} & \begin{pmatrix} -1 & 1 & 3 \\ 8 & 5 & 2 \\ 2 & 2 & 7 \end{pmatrix} \\ \begin{pmatrix} 1 & 5 & -2 \\ 3 & 1 & 1 \end{pmatrix} & \begin{pmatrix} \mathbf{35} & \mathbf{22} & \mathbf{-1} \\ \mathbf{7} & \mathbf{10} & \mathbf{18} \end{pmatrix} \end{array}$$

avec $C_{1,1} = 1*(-1)+5*8+(-2)*2= -1+40-4 =35$

Remarque : on peut calculer AB mais pas BA

Puissance de matrice :

- Elle n'est possible uniquement pour des **matrices carrées.**
- En effet pour effectuer un produit de matrices il faut que **le nombre de lignes de la deuxième matrice soit égale au nombre de colonnes de la première matrice**, or dans ce cas B=A et donc n=p.

Exemple : On peut calculer $A = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}^2$ mais pas $B = \cancel{\begin{pmatrix} 8 \\ 6 \end{pmatrix}^2}$

Paires de matrices commutantes :

- Dans le cas général, le produit de matrice AB est différent de BA. Il se peut même que l'un des produits n'existe pas. Dans l'exemple précédent, le produit BA n'est pas possible car le nombre de colonnes de B est différent du nombre de lignes de A.

  **Si AB = BA on dit que les matrices commutent** +++

- Aussi, pour toute matrice carrée, si $A^{-1}$ existe, alors $A^{-1}$ et A commutent toujours.

## III. Calcul matriciel pour LES ANALYSES MULTIVARIÉES (analyses factorielles) - modélisation en santé

Le terme **d'analyses factorielles** désigne un ensemble de techniques d'ajustement linéaire dont le but est de résumer l'essentiel de l'information contenue dans des gros tableaux de données (plusieurs dizaines, centaines d'individus observés selon un grand nombre de variables (supérieures à 100 par exemple)). Le procédé consiste à passer d'un espace de grandes dimensions à un espace plus petit (factorisation du tableau de données) avec une **perte d'information minimale et contrôlée.**

2 techniques principales d'analyse factorielle :

• L'analyse en composantes principales ou ACP : la plus ancienne des deux (1933) mais véritablement développée avec l'informatique, employée dans le cadre de variables quantitatives, homogènes ou pas

• L'analyse factorielle des correspondances ou AFC : (années 70), pour l'étude de tableaux de contingence (données qualitatives)

• D'un point de vue méthodologique, le fonctionnement des deux méthodes, est le même, on ne décrira ici que l'ACP

# A. ACP- INTÉRÊT ET ANALYSE DE DONNÉES

- L'intérêt de l'ACP est de :

- **Extraire** le maximum d'informations sous une forme simple et cohérente à partir d'un ensemble très important de données (description synthétique)

- Permettre de **mettre en évidence**
  - Des interrelations entre variables (redondance)
  - Des ressemblances ou oppositions entre individus (profils)

**Les résultats se présentent sous forme de combinaisons linéaires de variables différenciant les individus statistiques**

- Cas d'application :

- Uniquement sur des variables **quantitatives**
  - Elles peuvent être exprimées dans une même unité
    Exemple : le % de cas de Covid 19 dans les passages aux urgences
  - Elles peuvent être exprimées dans des unités différentes Exemple : la mort infantile, le revenu par habitant

Le tableau de données est constitué de n individus statistiques (unités spatiales, individus ...) caractérisés par p variables quantitatives. Ce tableau de données D constitue une matrice d'informations (n,p).

• **Chaque ligne** = vecteur ligne décrit un individu selon p variables

• **Chaque colonne** = vecteur colonne décrit un indicateur selon n

individus. Exemple :

| | Taille (m) | Masse (Kg) | IMC | Tour de tête (cm) | Age (années) | Moyenne au bac | ... |
|---|---|---|---|---|---|---|---|
| Patient 1 | 1.60 | 50 | 19.531 | 55 | 17 | 16 | ... |
| Patient 2 | 1.82 | 78 | 23.548 | 61 | 25 | 13 | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |

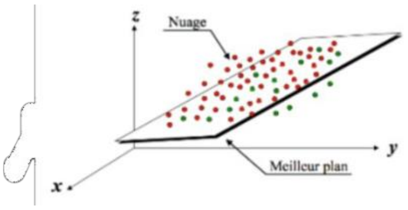
- Ainsi, le **but** de l'ACP est de prendre l'information contenue dans ce tableau de dimension importante et la représenter sous forme simplifiée.

# B. ACP- MÉTHODE

- L'ACP consiste à réduire la taille du nuage de points multi-dimensionnels en un nuage de points en 3-4 dimensions.

**Point tut'** : Lorsqu'on analyse **1 variable**, on a besoin que d'une dimension pour la représenter : c'est la représentation axiale. Pour analyser **2 variables**, on passe à 2 dimensions (sur un graph, on peut par exemple représenter la masse sur l'axe des abscisses et la taille sur l'axe des ordonnées). Quand on ajoute une **3ème variable**, on passe à 3 dimensions (cf image), on ajoute un nouvel axe et ainsi de suite.

**Au-delà de 3 variables, on va utiliser l'ACP qui permettra de condenser les informations de variables différentes dans les premiers axes**




- Pour cela on fait une **projection selon des axes** (axes factoriels ou facteurs). Ces axes sont des combinaisons linéaires de variables .

Fi=A1X1+A2X2+.....ApXp avec la plupart des Ai=0

- **Les coefficients Ai** permettent de mesurer l'intensité de la relation de chaque variable avec le facteur considéré (de max à nul selon les composantes), les coefficients Ai changent d'un facteur à l'autre. *Les coefficients étant les données que l'on trouve dans les matrices. Les coefficients Ai servent comme dans une régression multiple, à exprimer la part de chaque variable dans la constitution de l'axe.*
- **Les facteurs sont hiérarchisés** : l'axe 1 compte le maximum d'informations, c'est l'axe de plus grande dispersion du nuage de points, mais il laisse de côté des résidus. C'est le 2ième axe qui prend en compte le maximum d'informations résiduelles et ainsi de suite pour les axes suivants. Par construction, tous les axes (facteurs) sont **non corrélés**, ils forment donc des angles droits deux à deux.

# C. ACP- CALCUL DES AXES FACTORIELS

- L'idée est de transformer la matrice d'information en une matrice de projection des individus statistiques sur les axes..
- Si les données sont assez **homogènes**, on pratique un simple **centrage**.
- Si les données sont assez **hétérogènes**, le **centrer-réduire est obligatoire** (ramène la moyenne à 0 et l'écart-type à 1).
  - Centrer–réduire les données permet de gommer les effets taille et donne une ACP normée.

L'ACP normée consiste en :

- Variables centrées réduites
- Projections orthogonales
- Méthode des moindres carrés

**Point tut'** : Les axes (=facteurs) sont hiérarchisés : le premier axe contient le **maximum** d'informations, c'est l'axe sur lequel le nuage se déforme le moins possible en projection. On cherche ensuite le second axe sur lequel le nuage se déforme le moins après le premier axe (en projection), tout étant orthogonal (perpendiculaire) au premier axe. On réitère jusqu'à l'obtention de p axes. Le **meilleur axe** (premier axe) est celui où la dispersion du nuage de points est maximale, c'est-à-dire où **l'inertie du nuage est maximale**. On va voir la notion de « matrice inertie ».

## Détermination des axes factoriels :

La matrice de données D (individus variables) est une matrice à n lignes et p colonnes.
À cette matrice on associe une matrice D', la **matrice transposée** de D. En faisant le produit de D'.D on obtient une **matrice carrée**, **symétrique d'ordre p** appelée matrice d'inertie, notée T.
Toute l'ACP repose sur du **calcul matriciel**. Les axes sont définis par des vecteurs propres et valeurs propres :

- **Un vecteur propre est un vecteur tel que T.V=μV où μ est une valeur propre.**
  - Les valeurs possibles de μ sont déterminées en résolvant det (T-μI)=0
  - V a une seule colonne, et a autant de lignes que la matrice T a de colonne
  - Les vecteurs V sont normés

Exemple : Soit T = $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$
Pour trouver la valeur propre μ, on résout le système T.V=μV
$T = \mu I$

On a donc $T - \mu I = 0$, donc $T - \mu I = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - \mu \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-\mu & 2 \\ 2 & 1-\mu \end{pmatrix} = 0$

*En effet, quand on divise par V des 2 côtés de l'égalité, comme c'est du calcul matriciel on ne peut pas écrire T= μ, on est obligé de garder une forme matricielle à gauche comme à droite, donc on multiplie par I, ici la matrice identité d'ordre 2.*

*Pour trouver une solution à ce type d'équation, on cherche à ce que le déterminant de la matrice $T - \mu I$ soit égal à 0. Ainsi :*

Det $(T - \mu I) = (1-\mu)^2 - 4 = (-1-\mu)(3-\mu) = 0$
La résolution de cette équation donne μ=-1 et μ=3.
On résout alors les systèmes T.V=-V et T.V=3V (le prof ne détaille que le 2ème système).

$$T.V = 3V \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \times V = 3V$$
$$\Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} V1 \\ V2 \end{pmatrix} = 3 \times \begin{pmatrix} V1 \\ V2 \end{pmatrix}$$
$$\Leftrightarrow \begin{pmatrix} V1 + 2V2 \\ 2V1 + V2 \end{pmatrix} = \begin{pmatrix} 3V1 \\ 3V2 \end{pmatrix}$$
$$\Leftrightarrow \{V1+2V2 = 3V1 \text{ et } 2V1 + V2 = 3V2$$
$$\Leftrightarrow V1=V2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \textit{ car les vecteurs sont normés}$$

De même l'autre équation mène à $V3 = V4 = \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$

On remarque que les deux vecteurs obtenus sont bien orthogonaux (leur produit scalaire est nul : les axes sont donc bien décorrélés).

## D. ACP – part d'explication des axes factoriels

- La **part d'explication** d'un axe est donné par la formule suivante :

$$\mu_i\,\% = \frac{\mu_i}{\sum \mu_i} \times 100$$

- Dans le cas à n dimensions, deux situations peuvent se produire :
    - Soit l'histogramme des valeurs propres est **assez droit**, le nuage de points est plutôt arrondi sans axe d'allongement véritablement marqué : on peut en déduire que **les interrelations entre les variables sont sans doute faibles** et qu'il ne se dégage pas de combinaisons simples de l'ensemble des données.
    - Soit l'histogramme des valeurs propres est **assez concentré**, les valeurs propres sont très différenciées et on perçoit l'existence de deux axes d'allongement très marqués, on peut s'attendre alors à ce qu'il ressorte une **structure de différenciation forte** (profils marqués)

# E. ACP- Interprétation d'une analyse factorielle

Plusieurs données doivent être analysées :

- **Les coordonnées sur les axes factoriels** : elles donnent la position des individus par rapport aux axes factoriels. On peut alors mettre en évidence des oppositions entre groupe d'individus par rapport aux combinaisons de variables définies par les axes.
- **La qualité de représentation des individus sur les axes factoriels** : deux points distincts peuvent avoir la même projection sur l'axe factoriel, mais l'un sera mieux représenté que l'autre (angle plus petit).
    - Exemple : les variables 1 et 3 sont bien corrélées à l'axe k, elles forment un angle aigu avec l'axe k (bonne saturation sur l'axe), elles sont bien représentées. Tandis que la variable 2 est orthogonale à l'axe k (saturation nulle), elle n'a pas une bonne qualité de représentation.
- **La contribution des individus dans la formation de l'axe** : les individus contribuent plus ou moins à déterminer la direction des différents axes d'allongement du nuage. Elle est mesurée par la part de l'individu dans la variance.
- **La part de l'individu dans l'inertie totale du nuage = INR** : elle est proportionnelle à sa distance au centre de gravité (G). Elle donne une idée de la spécificité de l'individu par rapport à la moyenne.

C'est la fin pour ce cours !! Encore une fois avec de l'entraînement ça va aller, concentrez vous d'abord sur la partie calcul matriciel qui est la partie qui tombe +++ à l'examen.
Pour la partie ACP essayez de retenir les quelques infos importantes mais ce sont des questions simples de cours qui tombent.
On s'entraîne ++ avec les QCMS.
Courage 😸😸 si vous avez des questions ou autre n'hésitez surtout pas à m'envoyer un message sur messenger (Lozato Paolina) ou sur le forum/ Discord (Neuro'lina).
🤠