Iwatarax

Matrices

Matrices

Petit mot

Bon, bon, bon... Bienvenue face à ce cours autant aimé par les uns que détesté par les autres : les ✨MATRICES✨

Pour la plupart d'entre vous, la notion de matrice vous est encore inconnue (et c'est AUKAY), mais promis je vais faire de mon mieux pour changer ça et faire en sorte que tout soit le plus clair possible.

Pour résumer, on parlera aujourd'hui de modélisation en santé, de calcul matriciel et d'analyse factorielle (ne vous inquiétez pas, on va y venir).

Modélisation et algèbre linéaire

Modélisation en santé

La **modélisation**, en particulier dans le domaine de la santé, permet de reproduire informatiquement une situation dans le but de tester des scénarios ou des configurations particulières.

Dans un contexte de Big Data, la manière la plus "simple" de traiter les données (massives et diverses) est de les structurer au sein de GRANDS tableaux : les **matrices**

Sur ces matrices, on peut effectuer toutes sortes d'opérations mathématiques, ce qui est très précieux dans le domaine des statistiques multivariées *(ex : effet d'une hausse du prix des cigarettes sur le système de santé)*.

Bases de l'algèbre linéaire

L'**algèbre linéaire** correspond à la branche des mathématiques qui étudie les transformations linéaires et les espaces vectoriels.

Un **espace vectoriel** est une structure stable par addition de vecteurs et par multiplication par un scalaire. Autrement dit, si on additionne des vecteurs ou si on les multiplie par un nombre, le résultat appartiendra à l'espace de départ.

Calcul matriciel

Définition

Une matrice, notée *M(n,p)*, correspond à un tableau de nombres qui se compose de *n* lignes et *p* colonnes.

On distingue 2 cas :
- *p = 1* : on parle de "**univarié**" (matrice colonne)
- *p ≥ 2* : on parle de "**multivarié**"

$$\text{Exemples : } A = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \text{ est univarié et } B = \begin{pmatrix} 2 & 1 & 4 \\ 3 & 9 & 5 \end{pmatrix} \text{ est multivarié}$$

Matrices particulières

1. Matrice carrée

Une matrice est dite "**carrée**" d'ordre *n* lorsqu'elle possède autant de lignes *(n)* que de colonnes *(n)*.

$$\text{Exemple : } A = \begin{pmatrix} 1 & 4 \\ 3 & 7 \end{pmatrix} \text{ est une matrice carrée d'ordre 2}$$

2. Matrice identité

Une matrice est dite "**identité**" d'ordre *n* lorsqu'elle est carrée d'ordre *n*, que tous ses coefficients diagonaux valent 1 et que tous les autres valent 0.

$$\text{Exemple : } A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ est la matrice identité d'ordre 2}$$

Opérations de matrices

1. Transposée d'une matrice

La **transposée** d'une matrice *A(n,p)*, notée $^{T}A(p,n)$, correspond à une matrice à *p* lignes et *n* colonnes.

La transposée d'une matrice, qui existe **TOUJOURS**, revient donc à présenter l'information qui est en colonnes en lignes, et inversement.

$$\text{Exemple : la transposée de } A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \text{ est } {}^{T}A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

On distingue 2 cas particuliers :

- ${}^{T}A = A$: la matrice A est dite "**symétrique**"
- ${}^{T}A = -A$: la matrice A est dite "**antisymétrique**"

Exemples : $A = \begin{pmatrix} 1 & 3 \\ 3 & 5 \end{pmatrix}$ *est symétrique et* $A = \begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}$ *est antisymétrique*

Le produit ${}^{T}A \times A$ donne **TOUJOURS** comme résultat une matrice carrée d'ordre p ayant comme particularité d'être symétrique (par rapport à la diagonale).

Exemple : soit $A = \begin{pmatrix} 1 & 3 \\ 5 & 2 \\ 0 & 4 \end{pmatrix}$, *on a* ${}^{T}A \times A = \begin{pmatrix} 26 & 13 \\ 13 & 29 \end{pmatrix}$

PS : ne vous inquiétez pas si vous ne comprenez pas d'où sort ce résultat, on voit les produits de matrices un peu plus loin

2. Déterminant d'une matrice

Le **déterminant** d'une matrice (carrée) est une valeur particulière qui fournit des informations importantes sur cette matrice.

PS : il sera notamment utile pour déterminer si une matrice est inversible

De manière générale, il vous sera demandé d'être capable de calculer le déterminant de matrices d'ordres 2 et 3.

Pour un ordre 2 : $\boxed{det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc}$

Exemple :
$$\begin{aligned} det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} &= 1 \times 4 - 2 \times 3 \\ &= 4 - 6 \\ &= -2 \end{aligned}$$

Pour un ordre 3 : $\boxed{det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \times det\begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \times det\begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \times det\begin{pmatrix} d & e \\ g & h \end{pmatrix}}$

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

PS : si vous comprenez la logique du calcul, c'est parfait !

3. Inverse d'une matrice

L'**inverse** d'une matrice, qui existe **UNIQUEMENT** pour les matrices carrées dont le déterminant est non nul, est défini comme A^{-1} tel que : $A \times A^{-1} = I$

PS : I correspond à la matrice identité de même ordre que la matrice étudiée

Pour un ordre 2 : $\boxed{A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow A^{-1} = \frac{1}{det(A)} \times \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$

Exemple : soit $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ *, on a* $A^{-1} = -\frac{1}{2} \times \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1,5 & -0,5 \end{pmatrix}$

On note que la transposée d'une matrice inversible est **TOUJOURS** inversible.

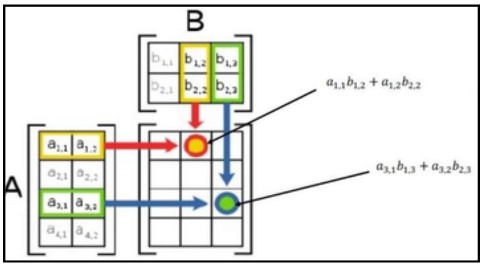
4. Produit de matrices

Le **produit** de 2 matrices *A* et *B*, noté *A×B*, existe **UNIQUEMENT** lorsque le nombre de colonnes de *A* est égal au nombre de lignes de *B*.

Ainsi, le produit de *A(n,p)* et de *B(p,m)* donne une matrice de la forme *C(n,m)*.

Pour calculer le produit *A×B*, il faut suivre une méthode particulière (cf. schéma).

PS : j'avoue que c'est pas le schéma le plus clair du monde, mais ça reste celui du cours du prof donc je vous le mets



Exemple d'un calcul de produit

Soient *A* et *B* deux matrices carrées d'ordre 2 (cas simple) :

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}$$

Le produit *A×B* se calcule de la manière suivante :

$$A \times B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \times \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \times 5 + 2 \times 7 & 1 \times 6 + 2 \times 8 \\ 3 \times 5 + 4 \times 7 & 3 \times 6 + 4 \times 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

PS : pour ceux qui ne verraient pas la logique, n'hésitez pas à jeter un coup d'œil à mon diapo Bonus sur le forum :)

Dans le cas général, la multiplication de matrices n'est pas **commutative**, autrement dit le résultat du produit $A×B$ n'est pas nécessairement (voire peu souvent) le même que celui du produit $B×A$.

PS : il est même possible que l'un des 2 produits n'existe tout simplement pas.

Lorsque $A×B = B×A$, on dit que les matrices A et B commutent : on parle de **paire de matrices commutantes**

On note que pour toute matrice carrée inversible *(rappel : det(A) ≠ 0)*, le produit $A×A^{-1}$ est nécessairement commutatif, autrement dit $A×A^{-1} = A^{-1}×A$.

Autre remarque intéressante, le produit de 2 matrices peut donner une matrice nulle **MÊME SI** aucune des 2 matrices n'est nulle.

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

5. Puissance d'une matrice

La **puissance** d'une matrice existe **UNIQUEMENT** pour des matrices carrées.

En effet, pour effectuer un produit de matrices $A×B$, il faut que le nombre de colonnes de A soit égal au nombre de lignes de B *(#rappel)*.

Or, dans le cas où $B = A$ *(ex : $A^2 = A×A$)*, il est nécessaire d'avoir la relation $n = p$ pour que le calcul soit réalisable, autrement dit la matrice A doit être carrée.

Exemple : on peut calculer $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix}$, *mais pas* $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^2$

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

Exemple : $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ *est nilpotente d'ordre 3 car* $A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ *et* $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Vous venez d'atteindre la fin de la partie sur le calcul matriciel, BRAVO !!

Avant de passer à la suite, qui traite du concept d'analyse factorielle, prenez une petite pause si vous en sentez l'envie.

Analyse factorielle

Généralités

Les **analyses factorielles** correspondent à un ensemble de techniques d'ajustement linéaire dont le but est de résumer l'essentiel de l'information contenue dans de gros tableaux de données.

Concrètement, ces analyses permettent de passer d'un espace de grandes dimensions (gros tableaux) à un espace plus petit (factorisation du tableau) tout en ayant une perte d'information minimale et contrôlée.

On distingue 2 techniques principales :

- **Analyse en composantes principales** (ACP) : technique plus ancienne (1933) employée dans le cadre de variables quantitatives
- **Analyse factorielle des correspondances** (AFC) : technique plus récente (1970s) employée pour les tableaux de contingence (variables qualitatives)

Le fonctionnement de ces 2 méthodes étant sensiblement le même, on ne s'intéressera dans ce cours qu'à l'ACP.

Intérêt et analyse de données

L'**ACP** permet d'extraire un maximum d'informations sous une forme simple et cohérente à partir d'une importante quantité de données.

Ses résultats se présentant sous forme de combinaisons linéaires de variables différenciant les individus statistiques, cette technique sert également à mettre en évidence les interrelations entre les variables (redondance) et les ressemblances ou oppositions entre les individus (profils).

L'ACP s'applique **UNIQUEMENT** sur des variables quantitatives, qui peuvent aussi bien être exprimées dans une même unité que des unités différentes.

Les **données** étudiées dans le cadre d'une ACP sont rassemblées dans des tableaux de données qui se composent de *n* individus statistiques caractérisés par *p* variables quantitatives : on parle de **matrice d'informations** $D(n,p)$

Chaque ligne représente un vecteur ligne décrivant un individu selon *p* variables tandis que chaque colonne est un vecteur colonne décrivant une variable selon *n* individus.

Axes factoriels

1. Principe

L'ACP consiste à **réduire** la taille du nuage de points multi-dimensionnels (données issues de notre fameux "gros tableau") en un nuage de points en 3 ou 4 dimensions.

Pour ce faire, il faut réaliser une projection selon des axes dits "**factoriels**" (ou "**facteurs**") qui correspondent à des combinaisons linéaires de variables.

Une **combinaison linéaire** a pour forme : $F_i = A_1 \times X_1 + A_2 \times X_2 + ... + A_p \times X_p$

- X_i : variables
- A_i : coefficients

Le rôle des **coefficients** est de mesurer l'intensité de la relation de chaque variable avec l'axe factoriel considéré, ils prennent donc des valeurs différentes d'un facteur à un autre *(rappel : axe factoriel = facteur)*.

Les facteurs sont dits "**hiérarchisés**", c'est-à-dire que l'axe 1 compte le maximum d'informations (c'est lui qui comprend la plus grande dispersion du nuage de points), puis l'axe 2 compte le maximum d'informations résiduelles (celles laissées de côté par l'axe 1) et ainsi de suite pour les axes suivants.

Par construction, tous les axes (ou facteurs) sont "**non corrélés**", autrement dit ils forment des angles droits 2 à 2.

2. Calcul des axes

L'idée consiste à transformer la matrice d'informations en une matrice de projection des individus statistiques sur les axes.

On peut faire face à 2 situations :

- **Données homogènes** : on réalise un simple centrage
- **Données hétérogènes** : il faut OBLIGATOIREMENT centrer-réduire

Centrer-réduire, c'est quoi ce truc ?

Centrer, c'est soustraire à une variable sa moyenne, ce qui permet de ramener la moyenne à **0**.

Réduire, c'est diviser une variable par son écart-type, ce qui permet de ramener l'écart-type à **1**.

Le fait de centrer (données homogènes) ou de centrer-réduire (données hétérogènes) permet d'aboutir à une ACP dite "**normée**".

Plus précisément, une ACP normée consiste en des variables centrées-réduites (1), des projections orthogonales (2) et la méthode des moindres carrées (3).

3. Détermination des axes

Pour rappel, la matrice d'informations *D* correspond à une matrice à *n* lignes (individus) et *p* colonnes (variables).

Le produit ${}^{T}D \times D$ donne une matrice carrée d'ordre *p* ayant comme particularité d'être symétrique *(#rappel)* : il s'agit de la **matrice d'inertie**, notée *T*

Toute l'ACP repose sur du calcul matriciel : les axes sont définis par des vecteurs propres et valeurs propres

PS : un vecteur propre, noté V, correspond à un vecteur tel que T×V = μ×V avec μ une valeur propre

Exemple de calcul d'axes factoriels

Soit $T = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ une matrice d'inertie.

Pour déterminer la valeur propre *μ*, on résout le système : *T×V = μ×V*

$$\begin{aligned} T \times V = \mu \times V &\Leftrightarrow T = \mu \times I \\ &\Leftrightarrow T - \mu \times I = 0 \\ &\Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} \mu & 0 \\ 0 & \mu \end{pmatrix} = 0 \\ &\Leftrightarrow \begin{pmatrix} 1-\mu & 2 \\ 2 & 1-\mu \end{pmatrix} = 0 \\ &\Leftrightarrow (1-\mu)^2 - 4 = 0 \\ &\Leftrightarrow (-1-\mu) \times (3-\mu) = 0 \end{aligned}$$

On obtient 2 solutions :
- *μ = -1*
- *μ = 3*

On doit maintenant résoudre les systèmes : *T×V = -V* et *T×V = 3V*

$$\begin{aligned} T \times V_1 = -V_1 &\Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = -\begin{pmatrix} a \\ b \end{pmatrix} \\ &\Leftrightarrow \begin{pmatrix} a+2b \\ 2a+b \end{pmatrix} = \begin{pmatrix} -a \\ -b \end{pmatrix} \\ &\Leftrightarrow \begin{cases} a+2b = -a \\ 2a+b = -b \end{cases} \\ &\Leftrightarrow \begin{cases} a = -b \\ b = -a \end{cases} \end{aligned}$$

$$\begin{aligned} T \times V_2 = 3V_2 &\Leftrightarrow \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = 3\begin{pmatrix} c \\ d \end{pmatrix} \\ &\Leftrightarrow \begin{pmatrix} c+2d \\ 2c+d \end{pmatrix} = \begin{pmatrix} 3c \\ 3d \end{pmatrix} \\ &\Leftrightarrow \begin{cases} c+2d = 3c \\ 2c+d = 3d \end{cases} \\ &\Leftrightarrow \begin{cases} c = d \\ d = c \end{cases} \end{aligned}$$

On trouve 2 vecteurs orthogonaux : $V_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$ $V_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

Explications bonus

Comment on passe de $T \times V = \mu \times I$ à $T = \mu \times I$?

Déjà, on divise par *V* des 2 côtés (jusqu'ici rien de choquant). Cependant, on n'a pas le droit de laisser simplement *T* = *μ* car il faut conserver une forme matricielle à gauche **ET** à droite.

C'est pourquoi il est nécessaire de multiplier par la matrice identité (ici d'ordre 2 puisque *T* est d'ordre 2).

Pourquoi le vecteur propre *V* est de la forme $V = \begin{pmatrix} a \\ b \end{pmatrix}$?

Un vecteur propre possède une unique colonne et autant de lignes que la matrice d'inertie *T* a de colonnes.

Dans l'exemple traité, la matrice d'inertie *T(2,2)* est carrée d'ordre 2, ce qui induit que le vecteur propre soit de la forme *V(2,1)*.

D'où ça sort ces coordonnées de fou pour les vecteurs V_1 et V_2 ???

Il faut savoir que les vecteurs propres *V* sont **normés**, c'est-à-dire que leur norme (leur "longueur") est égale à **1**.

La norme d'un vecteur $V = \begin{pmatrix} a \\ b \end{pmatrix}$ se calcule via la formule : $\|V\| = \sqrt{a^2 + b^2}$

Pour que cette norme soit égale à 1 et que l'on respecte les conditions trouvées via la résolution des systèmes, il faut nécessairement que les vecteurs V_1 et V_2 aient les coordonnées données plus tôt.

PS : j'espère que ces explications vous aideront à mieux comprendre :)

3. Part d'explication des axes

La part d'explication d'un axe se calcule via la formule : $\mu_i\% = \frac{\mu_i}{\sum \mu_i} \times 100$

Dans le cas à *n* dimensions, on distingue 2 situations :

- **Histogramme des valeurs propres droit** : on ne retrouve pas d'axe d'allongement véritablement marqué, ce qui peut signifier que les interrelations entre variables sont faibles (pas de combinaisons simples de l'ensemble des données)
- **Histogramme des valeurs propres concentré** : on retrouve 2 axes d'allongement très marqués, ce qui peut signifier qu'une structure de différenciation forte ressort (profils marqués)

Interprétation

Pour interpréter une analyse factorielle, il faut étudier plusieurs éléments :

- **Coordonnées sur les axes** : cela permet de connaître la position des individus par rapport aux axes, ce qui peut servir à mettre en évidence des oppositions entre des groupes d'individus
- **Qualité de représentation des individus sur les axes** : il est possible que 2 points distincts aient la même projection sur un axe, toutefois il peut y en avoir un mieux représenté (angle plus petit)
- **Contribution des individus dans les axes** : les individus ont plus ou moins de contribution dans la détermination de la direction des axes (cette contribution se mesure par la part de l'individu dans la variance)
- **Part de l'individu dans l'inertie totale du nuage** (INR) : cette donnée, qui est proportionnelle à la distance au centre de gravité, donne une idée de la spécificité de l'individu par rapport à la moyenne

FIN