Module A

ALGEBRE MATRICIELLE

Présentation - Plan

ALGEBRE MATRICIELLE

Ce module occupe une place particulière dans le dispositif de formation à la statistique mis en place par le service de Formation Permanente et le département de Biométrie de l'INRA. C'est un préalable aux modules : **Modèle linéaire**, **Analyse des données multidimensionnelles** et **Analyse discriminante**. Il a pour but de familiariser les stagiaires avec le concept matriciel et les concepts géométriques associés.

En effet, l'algèbre matricielle présente l'intérêt d'un formalisme simple très utilisé en statistique. De plus, elle permet de nombreuses représentations géométriques dans la mesure où on ne la considère pas comme une simple technique de calcul. Ces qualités sont particulièrement intéressantes en statistique puisqu'elles permettent :

- D'unifier la présentation des théories statistiques et d'automatiser les calculs.
- De rendre plus concrets les résumés statistiques par le truchement des représentations graphiques.

Ce dernier point est d'importance puisqu'il touche au domaine essentiel de l'interprétation. Dans ces conditions, on ne peut pas dissocier l'algèbre matricielle de la théorie géométrique qui la sous-tend.

Le contenu de ce module est :

- Un rappel des définitions élémentaires sur les espaces vectoriels et l'explicitation du concept de géométrie qui est la conséquence de la définition d'une distance qui dérive elle-même de la définition d'un produit scalaire.
- L'utilisation de la théorie des espaces vectoriels pour représenter géométriquement les résumés statistiques et probabilistes couramment utilisés et l'illustration par des exemples des principes géométriques présentés.
- Le calcul matriciel proprement dit et la signification géométrique des "objets matriciels" (vecteur, matrice, forme quadratique, valeur propre, ...). Enfin un exemple qui utilise les résultats d'algèbre matricielle qui ont été exposés pour décrire, le plus complètement possible, un ensemble de données.

Le contenu de la malette formation a été conçu par le groupe Animateur-Formateurs qui a tenu compte des critiques et commentaires faits par S. Junca-Holmes et

E. de Turckheim du Département de Biométrie de l'INRA et par G. Philippeau du service statistique de L'ITCF. Ces documents ont été mis en forme et encodés à la station de Biométrie et d'Intelligence Artificielle de Toulouse par R. Costes, M. Schnurrenberger et M.T. Tabes. Il contient :

- Les transparents.
- Le document-formateur comprenant :
  - Un volume "Algèbre matricielle" dans lequel le contenu du module A est exposé en détail.
  - Un volume "Transparents" en version réduite accompagnés de leurs commentaires.
- Le document-stagiaire qui est une version réduite des transparents à raison de deux transparents par feuille.
- Un questionnaire de "positionnement" des connaissances pour informer le formateur sur le niveau des stagiaires.
- Un questionnaire d'évaluation de la session de formation.
- Un document de présentation comprenant cette introduction et les tables des matières des deux volumes cités.

Compte tenu de la place particulière de ce module dans le plan de formation à la statistique il peut être considéré comme :

- Un module "à part entière" et dans ces conditions faire l'objet de sessions spécifiques de formation. Pratiquement, une session complète de ce module peut être faite en deux jours.
- Une source de connaissances à actualiser avant d'entreprendre des formations sur l'un des thèmes classiques de la statistique inférentielle ou descriptive.

Le contenu du volume "Algèbre matricielle" est plus large que celui du volume "Transparents" car le premier s'adresse aux formateurs et le second aux stagiaires. Cependant, les stagiaires peuvent être demandeurs d'exposés supplémentaires par rapport au programme qui leur est proposé et qui est matérialisé par le volume "Transparents". C'est pour cette raison que nous donnons dans ce qui suit le détail des points abordés dans chacun des volumes.

Le contenu du volume "Algèbre Matricielle"

1 Espace vectoriel - Géométrie

1.1 Quelques représentations géométriques

1.1.1 Egalité de deux vecteurs

1.1.2 Addition de deux vecteurs

1.1.3 Différence de deux vecteurs

1.1.4 Multiplication par un scalaire

1.1.5 Longueur, distance, angle, produit scalaire

1.2 Notion d'espace vectoriel

1.3 Projections

1.4 Géométrie analytique

2 Géométrie - Statistique - Probabilité

2.1 Statistique et géométrie

2.1.1 Résumer des observations

2.1.2 Mesurer des liaisons

2.1.3 Expliquer ou prédire des variables

2.1.4 Comparer des populations

2.2 Probabilité et géométrie

2.2.1 Espérance mathématique

2.2.2 Variance

2.2.3 Corrélation

2.2.4 Espérance conditionnelle

2.2.5 Variance et variance conditionnelle

2.2.6 Exemple

2.3 Compléments bibliographiques

2.4 Tableau résumé

3 Formalisation des concepts

3.1 Espace vectoriel

3.2 Sous espace vectoriel

3.3 Combinaison linéaire

3.4 Indépendance linéaire

3.5 Base d'un espace vectoriel

3.6 Le langage géométrique

3.7 Sous espace affine

3.8 Décomposition en somme directe

3.9 Exemples

3.9.1 Exemple 1 : représentations dans l'espace des variables

3.9.2 Exemple 2 : base d'un espace vectoriel

3.9.3 Exemple 3 : indépendance linéaire

3.9.4 Exemple 4 : application linéaire et représentation matricielle

4 Calcul matriciel

4.1 Définitions

4.2 Opérations élémentaires

4.2.1 Exemple 1 : écriture matricielle

4.2.2 Exemple 2 : Indépendance linéaire

4.3 Encore un peu de géométrie

4.4 Valeurs et vecteurs propres

4.4.1 Définition

4.4.2 Résultats généraux

4.4.3 Exemple 3 : recherche de valeurs et de vecteurs propres

4.4.4 Exemple 4 : vecteurs propres et droites de régression

4.5 Projecteurs associés à une décomposition en somme directe

4.5.1 Exemple 5 : projecteurs associés à des décompositions en somme directe

5 Partitionnement - Différenciation - Extremum

5.1 Partitionnement

5.1.1 Opérations élémentaires

5.1.2 Exemple 1 : inverse et déterminant d'une matrice partitionnée

5.2 Différenciation

5.3 Extremum

5.3.1 Maximisation sous contrainte

5.3.2 Interprétation géométrique des valeurs et des vecteurs propres

5.3.3 Relations entre deux ellipsoïdes

5.3.4 A propos de la maximisation

5.4 Dérivée d'un déterminant

5.5 Quelques formules utiles

5.6 Exemple 2 : description d'un ensemble d'observations

5.6.1 Illustration géométrique

Le contenu du volume "Transparents"

PARTIE I : ESPACE VECTORIEL

1 ESPACE VECTORIEL

- Addition de 2 vecteurs
- Multiplication par un scalaire
- Propriétés de l'addition
- Propriétés de la multiplication par un scalaire
- Propriétés déduites des précédentes

1.1 Sous espace vectoriel

1.2 Combinaison linéaire

1.3 Indépendance linéaire

1.4 Base d'un espace vectoriel

1.5 Longueur, distance, angle, produit scalaire

- Intérêt du produit scalaire
- Propriétés du produit scalaire

1.6 Décomposition en somme directe

2 PROJECTIONS

3 COORDONNEES DANS $I\!R^3$

PARTIE II : GEOMETRIE ET STATISTIQUE

1 REPRESENTATION DES DONNEES

- Les espaces des individus et des variables

2 LES RESUMES STATISTIQUES

- Moyenne et variance

3 LES LIAISONS STATISTIQUES

- Exemple : la liaison pression - température
- Notion de corrélation

3 LES TESTS STATISTIQUES

- Comparaison de deux populations

4 RESUME

PARTIE III : CALCUL MATRICIEL

1 MATRICES

- Ranger des données
- Traiter des données
- Faire une transformation

2 MATRICES PARTICULIERES

- Matrice carrée
- Matrice nulle
- Matrice diagonale
- Matrice identité I
- Matrice J
- Matrice transposée : A' transposée de A
- Matrice symétrique

3 OPERATIONS MATRICIELLES

- Addition

- Multiplication par un scalaire
- Multiplication de deux matrices
- Transposée d'un produit de matrices
- Trace
- Matrice idempotente
- Matrice inverse
- Notion de rang
- Déterminant d'une matrice

PARTIE IV-A : INTRODUCTION

1 INTRODUCTION

2 PRESENTATION DES DONNEES IRIS

3 PRESENTATION GRAPHIQUE

- Variance
- Covariance
- Coefficient de corrélation

4 PRESENTATION MATRICIELLE

- Variance-covariance
- Inertie

5 REPRESENTATION GEOMETRIQUE DE LA DISPERSION

- Une variable
- Deux variables
- Exemples

PARTIE IV-B : VALEURS ET VECTEURS PROPRES

1 NOTION DE VALEUR ET DE VECTEUR PROPRE

- Symétrie par rapport à la *1ère* bissectrice
- Généralisation

2 PROPRIETES

3 EXERCICES

- Symétrie par rapport à la *1ère* bissectrice
- L'exemple iris *setosa*

PARTIE IV-C : MAXIMISATION-DERIVATION

1 RECHERCHE DES EXTREMA

- Exemple de la cartouche de chasse
- Extremum d'une fonction

2 DERIVATION

- Dérivée première
- Exemple
- Dérivée seconde

3 RECHERCHE DES EXTREMA

- Maximisation du quotient de 2 formes quadratiques
- Maximisation d'une forme quadratique

PARTIE IV-D : L'EXEMPLE DES IRIS

1 DESCRIPTION DE LA POPULATION : IRIS *setosa*

- Projections sur les axes X et Y
- Recherche de l'axe qui maximise la dispersion
- Résumé

2 DESCRIPTION DES POPULATIONS IRIS *setosa* ET IRIS *versicolor*

- Estimation des paramètres de position et de dispersion
- Première direction propre de chaque population
- Matrice d'inertie commune aux deux populations
- Première direction propre commune aux deux populations
- Ellipses d'inertie des populations
- Matrice d'inertie des points moyens des populations
- Direction propre de la matrice d'inertie des points moyens
- Positions relatives des ellipses d'inertie intra et inter populations
- Recherche de l'axe qui sépare au mieux les deux populations