Explication - TD3 (AFD)

Auteur: Badr TAJINI - DataMining & AI - ESIEE - 2024/2025

Contexte conflictuel

Maximiser les distances entre les groupes revient à maximiser les projections suivantes :

(école Anglo saxonne)

$$P_1 = \max_u \left(\frac{u^t B u}{u^t W u} \right)$$

ou

(école Française)

$$P_2 = \max_u \left(\frac{u^t B u}{u^t W u} \right)$$

Rassurez-vous : il n'y a aucune contradiction. Ces deux « écoles » (anglo-saxonne vs. « à la française ») aboutissent au **même critère** mathématique, à savoir :

$$\max_u \quad \frac{u^T B\, u}{u^T W\, u}.$$

Pourquoi, alors, de nombreuses références les présentent-elles comme deux approches ? En pratique, la différence est surtout **conceptuelle** (dans la manière de motiver ou de dériver la formule) plutôt que **numérique** (la solution finale est la même). Voici les points qui clarifient ce paradoxe apparent.

---

1. Deux chemins, même destination

Approche anglo-saxonne

- Souvent introduite dans le contexte de l'**analyse discriminante** ou des **variantes canoniques** (Fisher's Linear Discriminant).
- Dérivation typique : « Nous voulons trouver une direction (u) qui maximise le ratio "entre-groupes" / "intra-groupes". »
- On résout un problème d'**autovaleurs généralisées** ($B\,u = \lambda\, W\, u$).

Approche « à la française »

- Souvent cadrée dans le contexte de l'**analyse des correspondances**, de l'**analyse factorielle discriminante**, ou du concept d'**inertie** en analyse de données.
- Dérivation typique : « Nous voulons trouver un sous-espace qui capte la plus grande différence entre les groupes par rapport à leur variabilité interne. »
- Après le développement de l'algèbre, on aboutit au **même** problème d'autovaleurs généralisées.

Ainsi, malgré des « histoires » différentes, l'**objectif** de maximiser

$$\frac{u^t B\, u}{u^t W\, u}$$

et l'**équation** finale restent les mêmes. On retrouve les mêmes vecteurs propres et valeurs propres qui maximisent ce ratio.

---

2. Différences mineures de notation et de contraintes

Même si les deux approches mènent au même ratio, on peut rencontrer :

1. **Des notations différentes pour (B) et (W).**
   - En anglais, on peut voir (S_B) et (S_W) (pour « between-class scatter » et « within-class scatter »).

   - Les références françaises parlent souvent d'**inertie**, de **covariance**, etc. Au fond, cela revient au même.
2. **Des contraintes légèrement différentes sur (u).**
   - Certains auteurs imposent ($\|u\| = 1$) (contrainte de norme), d'autres imposent ($u^t W\, u = 1$), etc.

   - Ces différences ne changent pas le **jeu de solutions** : elles n'affectent que la formulation algébrique ou le mode de normalisation.

En bref, le **même** problème d'autovaleurs généralisées peut sembler différent selon la normalisation ou le centrage des données. Les **vecteurs propres maximaux**, eux, restent identiques.

---

3. Même objectif, philosophie différente

Pourquoi ces deux « écoles » existent-elles ? Historiquement, elles s'appuient sur des **philosophies** un peu différentes de l'analyse de données :

- **Anglo-saxonne** : accent sur la **classification** statistique (discriminant, test d'hypothèse), par ex. « Cette direction sépare-t-elle le mieux les groupes ? ».
- **Française** : accent sur l'**analyse exploratoire** (analyse en composantes principales, analyse des correspondances), le concept d'**inertie**, la géométrie des nuages de points, etc.

Au bout du compte, une fois formulée la maximisation de

$$\frac{u^t B\, u}{u^t W\, u}$$

les deux approches aboutissent à **la même solution**. On peut voir cela comme un **unique problème** d'optimisation, raconté sous deux « histoires » ou traditions pédagogiques.

Conclusion

- Il n'y a **pas de conflit réel** : les deux méthodes résolvent le **même** problème mathématique.

- **La formule est bel et bien la même** ; toute divergence apparente tient à des différences de notation (B et W) ou de contraintes de normalisation sur (u).

- Les dérivations, bien qu'elles paraissent différentes, convergent vers le même problème d'**autovaleurs généralisées** et donnent les mêmes directions de séparation optimale.

En définitive, **les deux « écoles » ont raison** : elles ne sont que deux manières historiques (et pédagogiques) d'arriver à la même conclusion.