

Une introduction au test de dimension suffisante sur la prochaine valeur propre avec R

Pier-Olivier Caron¹ André Achim²

¹Université TÉLUQ ²Université du Québec à Montréal

Présentation au 46e congrès de la SQRP

Dans le cadre du symposium :
Analyse factorielle et modèle de Rasch : Une introduction
appliquée sur les meilleures pratiques contemporaines avec
le progiciel libre R

le 24 mai 2024

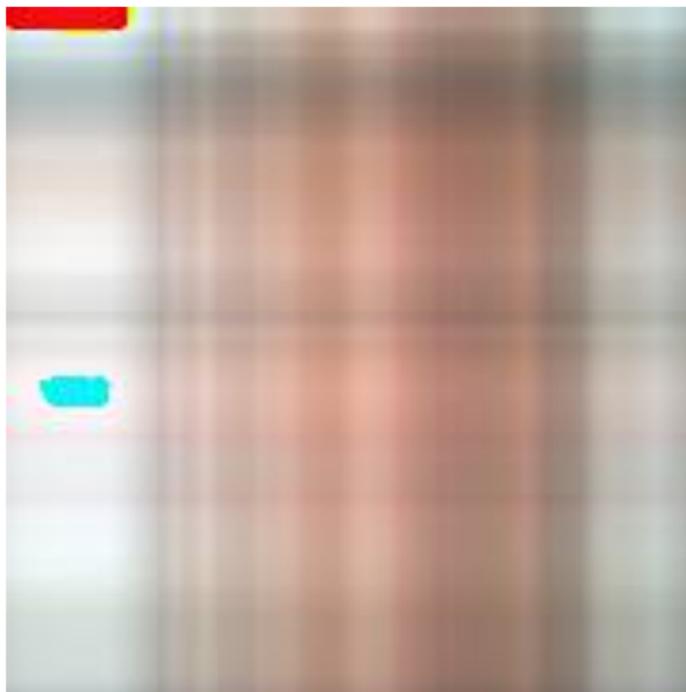
Introduction

Introduction

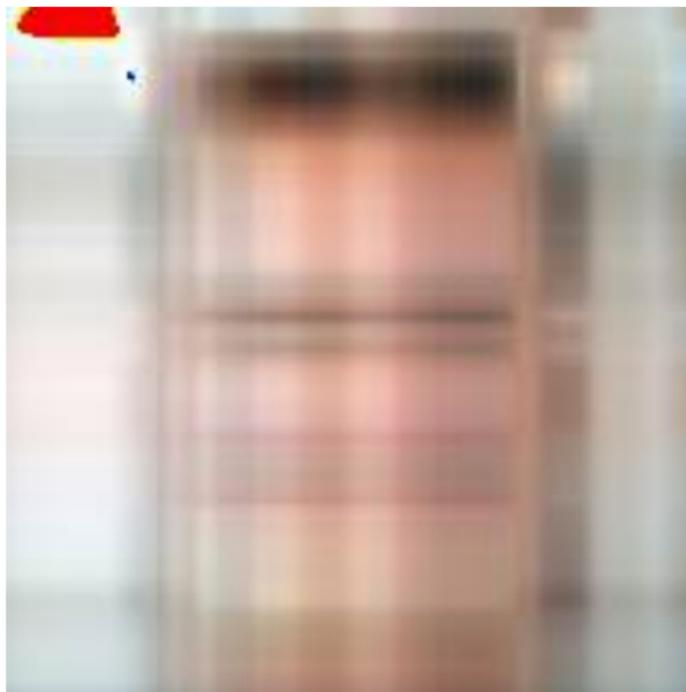
Déterminer les nombres de facteurs à retenir dans une analyse factorielle exploratoire (AFE) est l'un des problèmes méthodologiques toujours ouverts.

Qu'est-ce que réduire la dimensionalité?

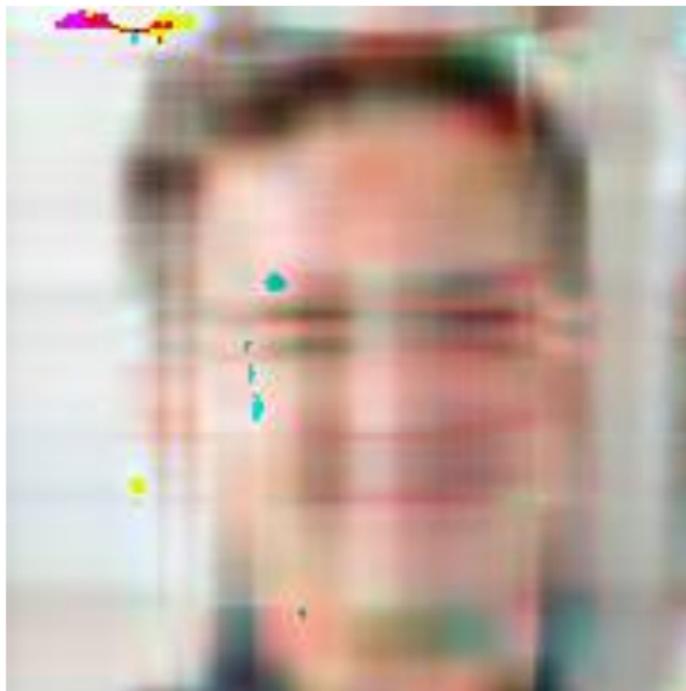
Facteur = 1



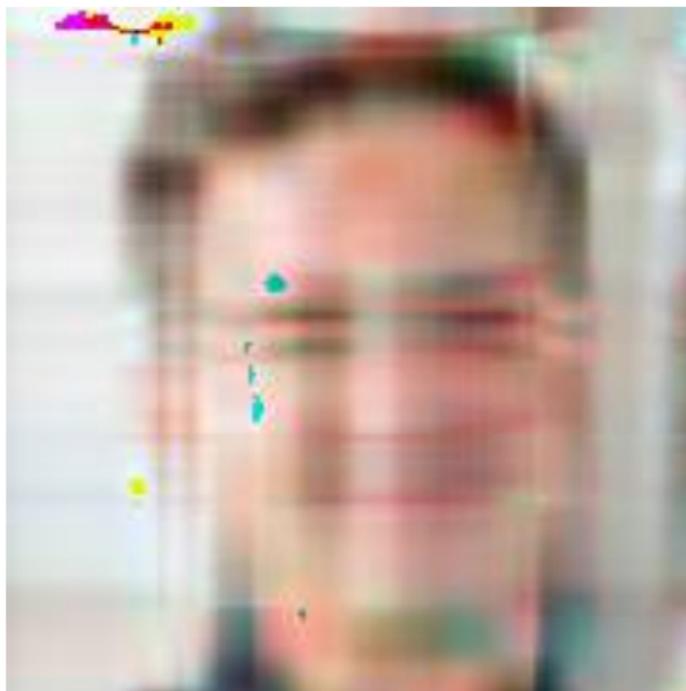
Facteurs = 2



Facteurs = 3



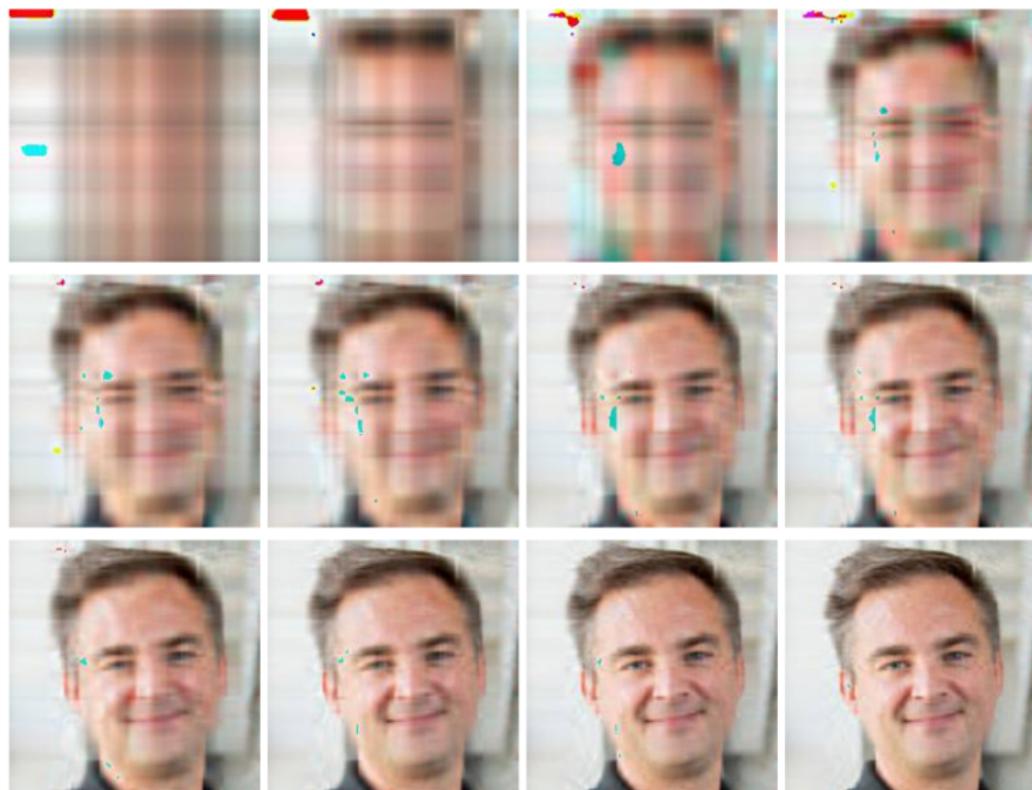
Facteurs = 5



Facteurs = 25



Facteurs incrémentés



Les règles d'arrêt

Les règles d'arrêt

L'AFE est agnostic au nombre de facteurs.

Il faut un critères pour arrêter de retirer des dimensions.

Il existe plusieurs méthodes et de nouvelles sont développées chaque année.

Le développement de NEST

Le critère de Kaiser (1960)

L'analyse parallèle de Horn (1965)

L'analyse parallèle de Green et al. (2012)

Le test de dimension suffisante sur la prochaine valeur propre (*Next Eigenvalue Sufficiency Tests*, NEST) (Achim, 2017, 2020, 2021).

NEST

NEST

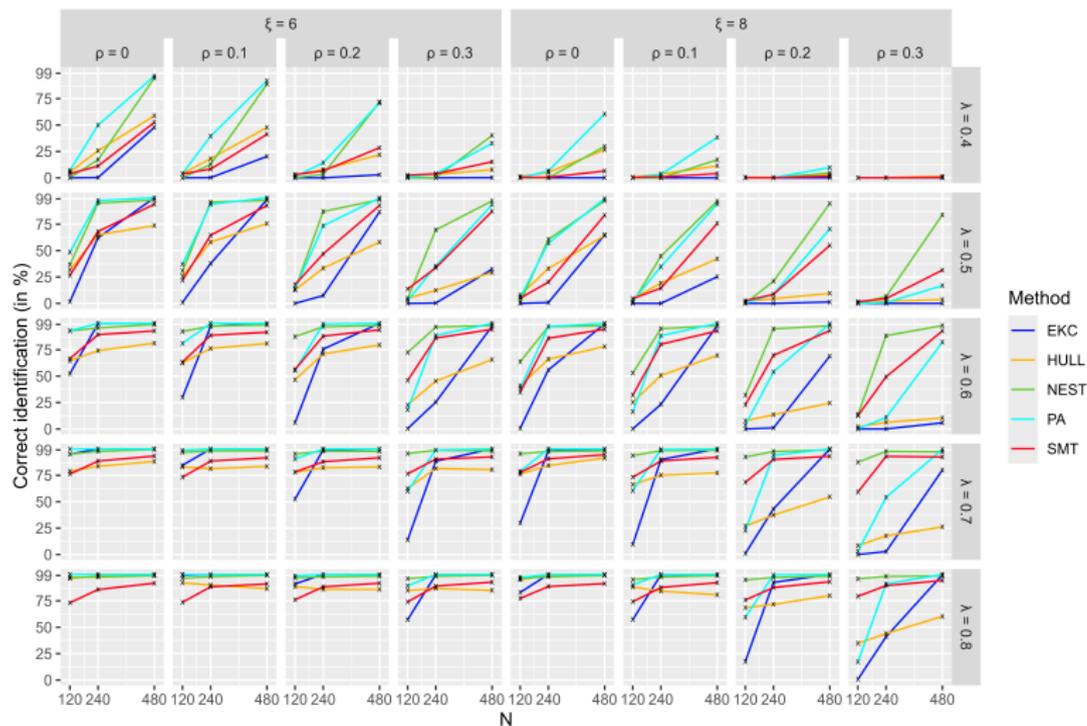
À chaque étape, on cherche un nombre k de dimensions qui soit nécessaires mais suffisantes pour ne laisser que du bruit dans les données.

Si k est suffisant, la valeur propre au rang $k + 1$ est quelconque (sous le 95^e percentile) parmi les valeurs propres de données simulées engendrées par le modèle à k facteurs.

Si k est trop petit, les corrélations entre variables ne sont pas adéquatement expliquées. Ce reste d'information partagée s'exprime dans une valeur propre supplémentaire au rang $k + 1$ (au-dessus du 95^e percentile).

La façon d'estimer ces modèles à k facteurs est critique. On ne pourrait pas prendre les k premières composantes principales.

Rendement de NEST (Caron, 2023)



Comment utiliser NEST avec R

```
# Installer NEST (version en développement)
remotes::install_github(repo = "quantmeth/Rnest")

# Bientôt
install.packages("Rnest")

# Importer jeu de données
jd <- foreign::read.spss("fichier.sav",
                        to.data.frame = TRUE,
                        use.value.labels = FALSE)

jd <- read::read_excel("fichier.xls")

jd <- read.csv2("fichier.csv")
```

Vérifications préliminaires

S'assurer que toutes les variables soient pertinentes (aucune orpheline)

S'assurer que les variables soient distribuées normalement

Vérifier la prépondérance des variables manquantes et les traiter

Sélectionner les variables de l'AFE

Rouler NEST

```
# Appeler NEST
library(Rnest)

# Rouler NEST avec un jeu de données
resultats <- nest(jd)

# Rouler NEST avec une matrice corrélation
mat.cor <- cor(jd)
resultats <- nest(mat.cor, n = nrow(jd))
```

La sortie

```
# La sortie  
resultats  
  
# Le sommaire  
summary(resultats)  
  
# Le graphique  
plot(resultats)
```

La sortie

```
At 95% confidence, Nest Eigenvalue Sufficiency Test (NEST)
suggests 2 factors.
```

Le sommaire

```
nest 0.0.0.1 ended normally
```

Estimator	ML
Number of model parameters	55
Resampling	1000
Sample size	404
Stopped at	3

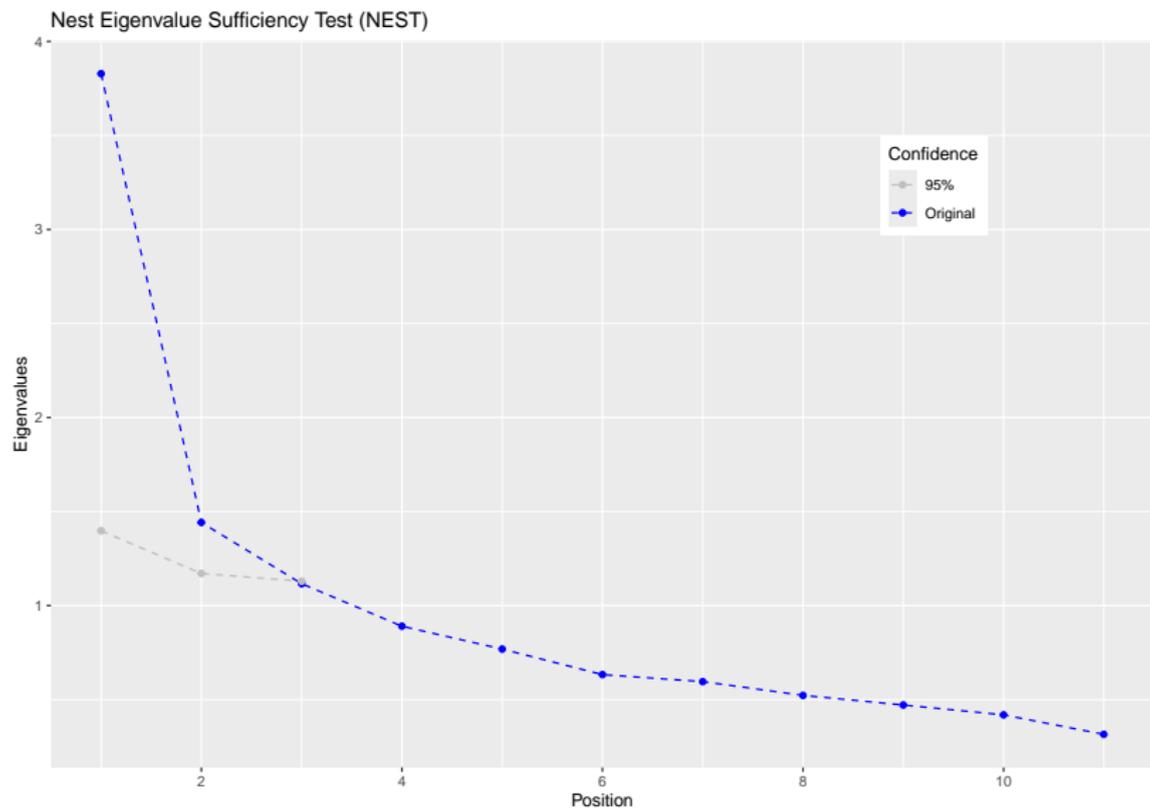
```
Probabilities of factors
```

Factor	Eigenvalue	Prob
F1	3.829	< .001
F2	1.442	< .001
F3	1.116	.077

```
At 95% confidence, Nest Eigenvalue Sufficiency  
Test (NEST) suggests 2 factors.
```

```
Try plot(nest()) to see a graphical  
representation of the results.
```

Le graphique



Quelques options

- `alpha` : permet de tester plusieurs seuils critiques
- `nrep` : permet de changer le nombre de rééchantillonnages
- `ggsave(plot(resultats), file = "figure_nest.png")`
permet de sauvegarder la figure

Comment utiliser NEST avec SPSS

```
BEGIN PROGRAM R.  
# S'assurer que R et Rnest sont installés dans SPSS  
# install.packages("Rnest")  
library(Rnest)  
data <- spssdata.GetDataFromSPSS()  
resultats <- nest(data)  
resultats  
summary(resultats)  
plot(resultats)  
END PROGRAM.
```

Comment utiliser NEST avec SPSS

Script disponible par Achim (2020)

<https://www.tqmp.org/RegularArticles/vol16-4/p213/>

Références I

- Achim, A. (2017). Testing the number of required dimensions in exploratory factor analysis. *The Quantitative Methods for Psychology*, 13(1), 64–74.
<https://doi.org/10.20982/tqmp.13.1.p064>
- Achim, A. (2020). Esprit et enjeux de l'analyse factorielle exploratoire. *The Quantitative Methods for Psychology*, 16(4), 213–247. <https://doi.org/10.20982/tqmp.16.4.p213>
- Achim, A. (2021). Determining the number of factors using parallel analysis and its recent variants: Comment on lim and jahng (2019). *Psychological Methods*, 26(1), 69–73.
<https://doi.org/10.1037/met0000269>

Références II

- Caron, P.-O. (2023). An evaluation of the nest eigenvalue sufficiency test (NEST). *Communication Presented at the 10th European Congress of Methodology of the European Association of Methodology*.
- Green, S. B., Levy, R., Thompson, M. S., Lu, M., & Lo, W.-J. (2012). A proposed solution to the problem with using completely random data to assess the number of factors with parallel analysis. *Educational and Psychological Measurement*, 72(3), 357–374. <https://doi.org/10.1177/0013164411422252>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>

Références III

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>