

TP3 : Bootstrap pour un modèle graphique

ENSTA

2021/2022

Adresse mail: mohammed.sedki@universite-paris-saclay.fr Page web: masedki.github.io

Les librairies

Nous allons avoir besoin des deux librairies

```
require("mvnfast")
require("igraph")
```

I. Générer le jeu de données

Soient X_1, \dots, X_n des vecteurs aléatoires à valeur dans \mathbb{R} . Soit Σ la matrice $d \times d$ de covariance de X_i . Le graphe des corrélations marginales place une arête entre (j, k) si $R(X_j, X_k) \neq 0$. Nous allons ici faire appel à une procédure bootstrap pour estimer le graphe des corrélations marginales. On suppose que Σ est de la forme suivante: $\Sigma_{jj} = 1, \Sigma_{j,k} = a$ si $|j - k| = 1$ et $\Sigma_{j,k} = 0$ sinon. Ici nous allons fixer $a = 1/4$.

On fixe la dimension de notre vecteur aléatoire à $d = 100$ et la taille du jeu de données à $n = 50$.

```
set.seed(7)
# Generate Data
n <- 50
d <- 100
sigma_mat <- toeplitz(c(1, 1/4, numeric(d-2)))
data <- rmvn(n, numeric(d), sigma_mat)
corr_data <- cor(data)
```

Calculer un intervalle de confiance à 95% par bootstrap pour la matrice de corrélation R . Cette procédure nous fournit un intervalle de confiance uniforme pour tous les éléments de R . Pour chaque couple de variables (j, k) , il existe une arête si l'intervalle de confiance exclut la valeur 0. Tracer le graphe. Répéter cette procédure pour différentes valeurs de a .