

Manifold Learning

Lab x002 : Géométrie des espaces de haute dimension

Jairo Cugliari

S1 2016–2017

1 Détection automatique de chiffres

1. Les données sont un sous ensemble du jeu de données MNIST Télécharger les données à utiliser.
2. Charger les données avec R. Observer que chaque ligne code un chiffre.
3. Écrire une fonction R qui permette transformer une ligne de la base en matrice. Ainsi, vous pourrez visualiser les chiffres à l'aide de la fonction `image`.
4. Faire un ACP sur le sous ensemble des données qui correspondent au chiffre 3.
5. Quels sont les modes de variation principaux du chiffre 3?

2 Distance minimale entre vecteurs aléatoires

Illustrez de manière numérique l'augmentation de la distance minimale entre deux points (tirés au hasard) quand la dimension D des points augmente.

3 Problème: Création de jeux de données artificiels

Simuler les jeux de données artificiels utilisés dans l'article Dimensionality Reduction: A Comparative Review (2009).

4 Exercices additionnels

ACP sur des ensembles convexes . Simuler un jeu de données contenant deux anneaux concentriques. Puis, obtenir la projection des points par ACP (linéaire). Représenter le nuage de points projetés, en colorant les points selon leur appartenance au noyau extérieur ou intérieur.