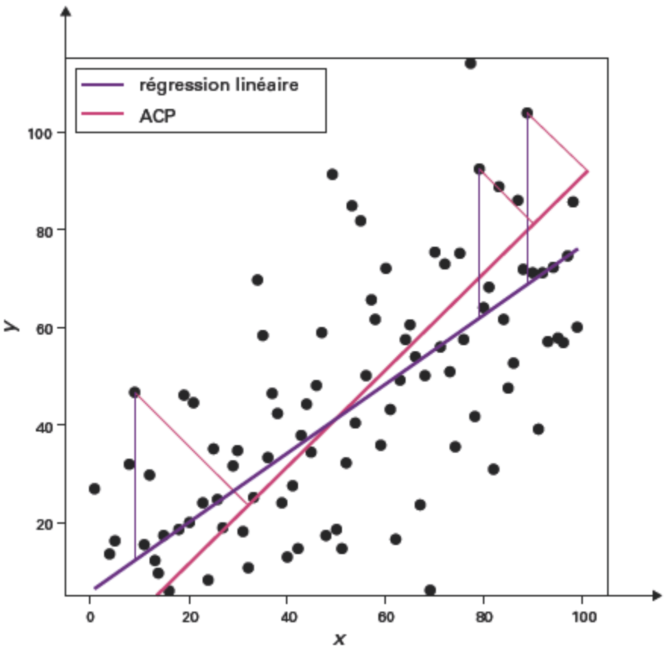

FIGURE K.13 – Figure : comparaison entre regression linéaire (apprentissage supervisé) et ACP (apprentissage non supervisé) appliquée à un nuage de points en 2 dimensions. On approxime le nuage par une droite, mais pour ACP (rose), on minimise les carrés des distances des points à la droite ; pour regression on minimise les carrés des erreurs de prédiction de y pour x fixé (segments verticaux). Sauvalle (2020)

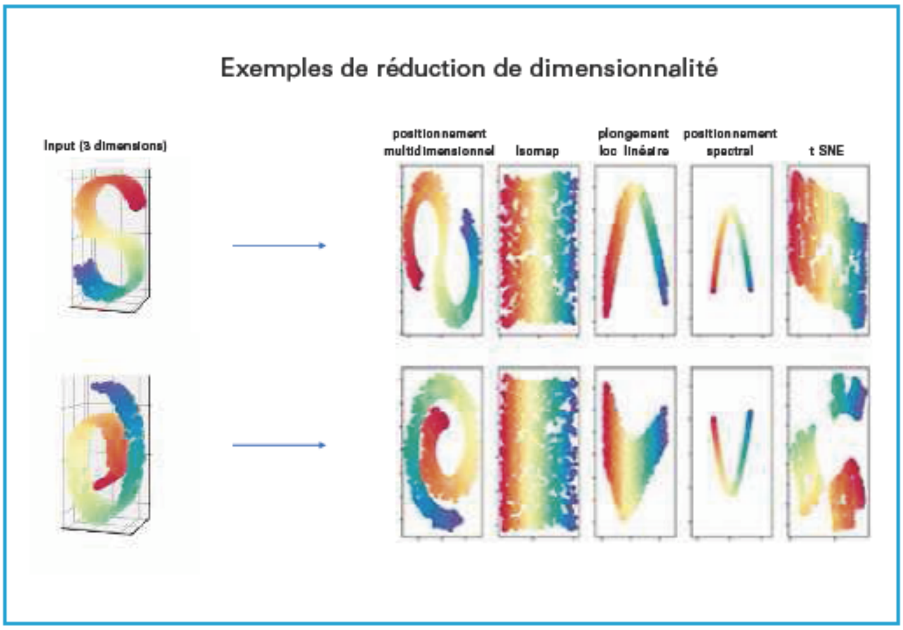

Figure 4 – Exemples de réductions de dimensionalité non linéaires utilisant positionnement multidimensionnel, Isomap, plongement localement linéaire et t-SNE (Images générées en utilisant scikit-learn)

FIGURE K.14 – Visualisation de distributions de données non linéaires et des résultats de méthodes de réduction de la dimensionalité (Sauvalle, 2020)

La figure 2.4 montre à gauche, un exemple de données 3D qui se trouvent dans un plan 2D et à droite les deux premières composantes principales sur ces données.

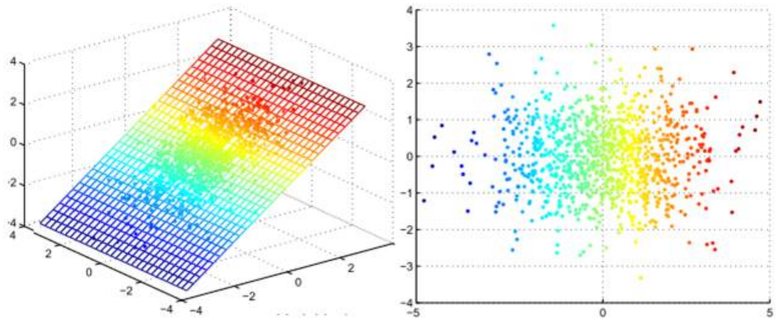

Figure 2.4 – ACP sur des données linéaires

Supposons que nous ayons un ensemble de données $X = \{x_1, x_2, .., x_M\}$ composé de M observations où chaque observation $x_i = (x_{i1}, x_{i2}, .., x_{iN})$ est composée de N caractéristiques. X est associé à une matrice de données A de taille $N \times M$ où chaque colonne représente une caractéristique. En pratique, le calcul de l'ACP pour la matrice X revient à réaliser les opérations ci-dessous afin de trouver les composantes principales :

1. Calculer le vecteur $\mu = (\mu_1, \mu_2, \cdots, \mu_m)^T$ qui représente le vecteur moyen où μ_i est la moyenne de la *ième* composante des données.
2. Calculer la matrice χ en soustrayant le vecteur moyen à toutes les colonnes de A dans le but d'obtenir des données centrées.
3. Calculer la matrice S (de taille $N \times N$) de covariance de χ avec ($S = \chi \cdot \chi^T$).
4. Calculer la matrice U (de taille $N \times N$) qui est composée des coordonnées des vecteurs propres $\vec{u_j}$ de S triés par ordre décroissant des modules des valeurs propres λ_j (la première colonne de U est le vecteur propre qui correspond à la plus grande valeur propre)
5. Garder les R premières colonnes de U pour former la matrice $\tilde{U} : N \times R$ qui représente les R premières composantes principales.

Figure K.15 – Algorithme d'ACP présenté par Chouaib (2011)