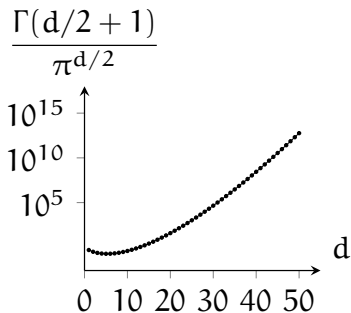
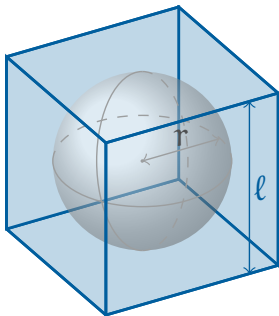


■ ANALYSE EN COMPOSANTES PRINCIPALES

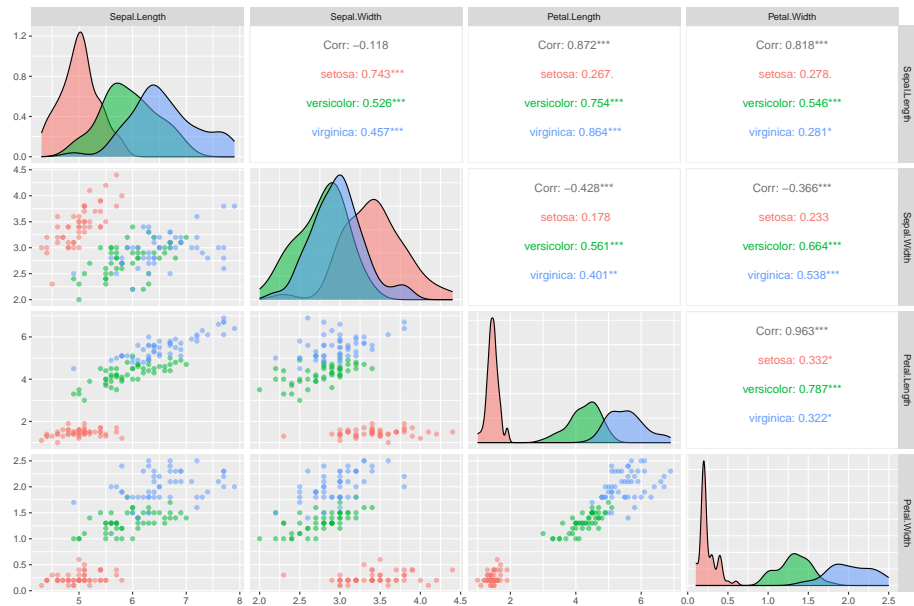
■ Fléau de la dimension

Dans \mathbb{R}^d \mathcal{S} boule de rayon r \mathcal{C} cube de côté $\ell = 2r$

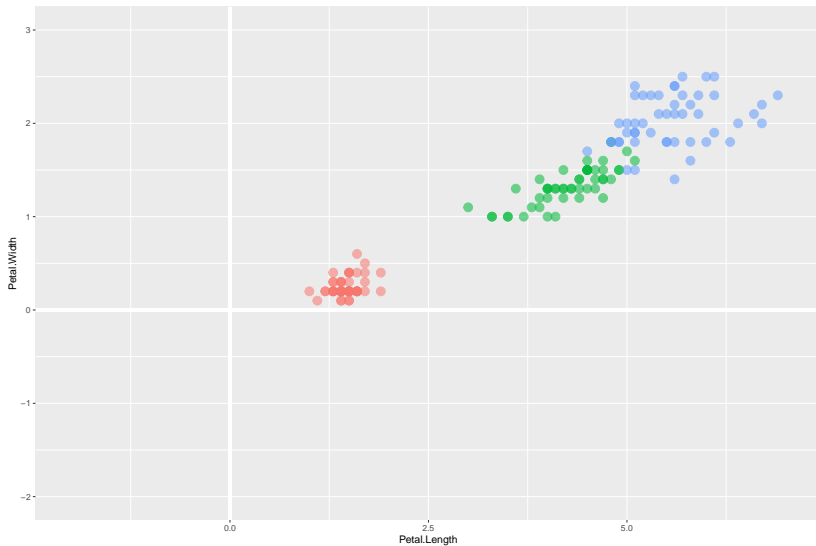
$$|\mathcal{C}| = \ell^d \quad |\mathcal{S}| = \frac{r^d \pi^{d/2}}{\Gamma(d/2 + 1)} \quad \frac{|\mathcal{C}|}{|\mathcal{S}|} = 2^d \frac{\Gamma(d/2 + 1)}{\pi^{d/2}}$$



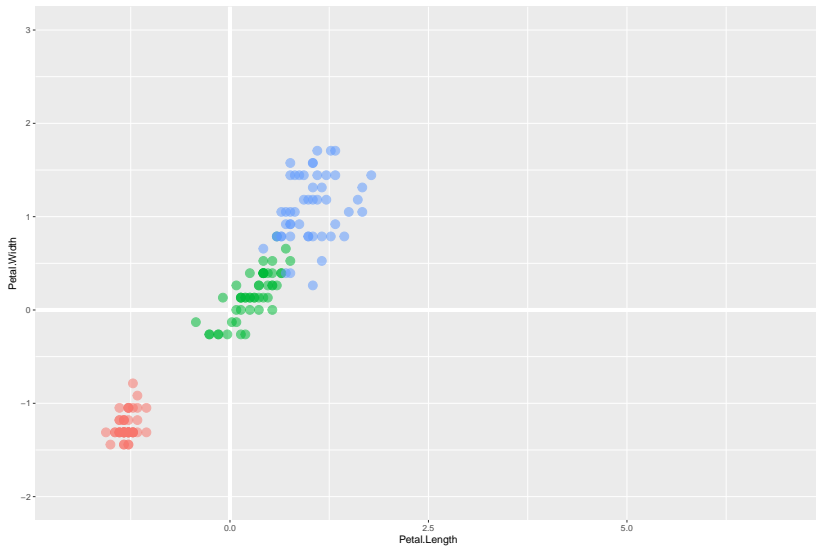
Le jeu de données Iris



Données brutes



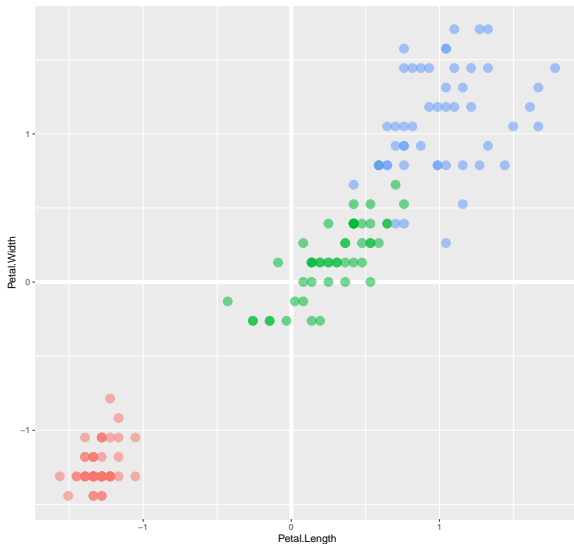
Données standardisées



■ Construction des axes principaux



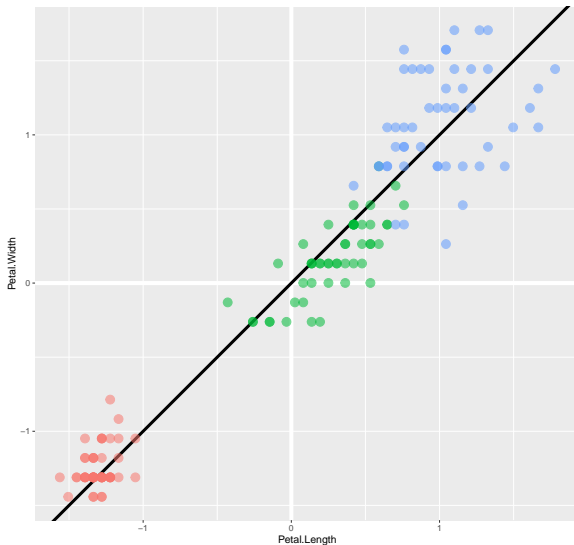
Données standardisées



■ Construction des axes principaux



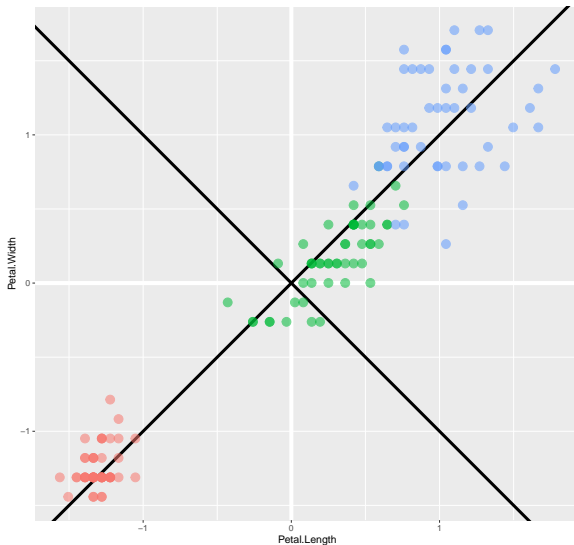
Premier axe principal



■ Construction des axes principaux



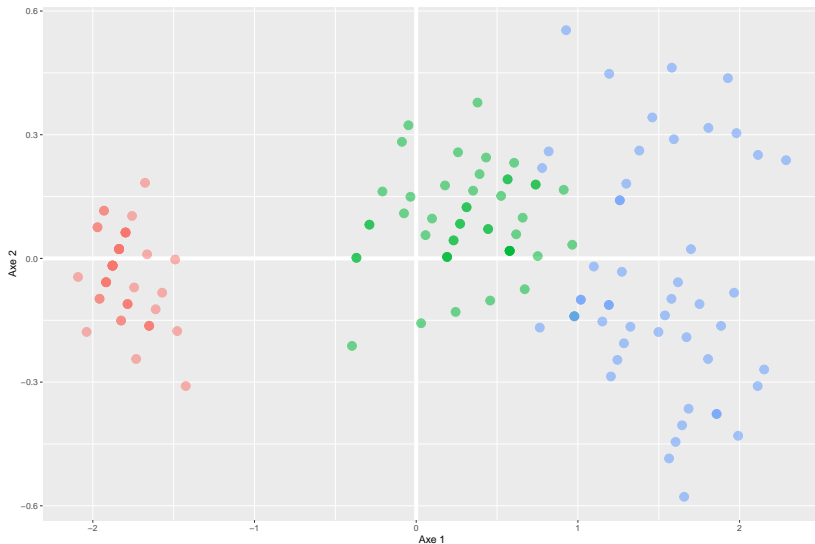
Deuxième axe principal



■ Construction des axes principaux



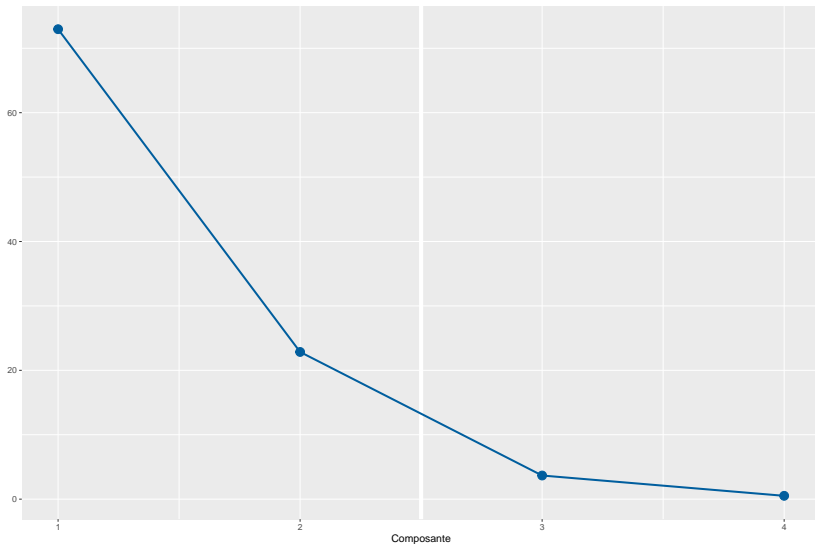
Données dans le nouveau repère



■ Choix du nombre d'axes



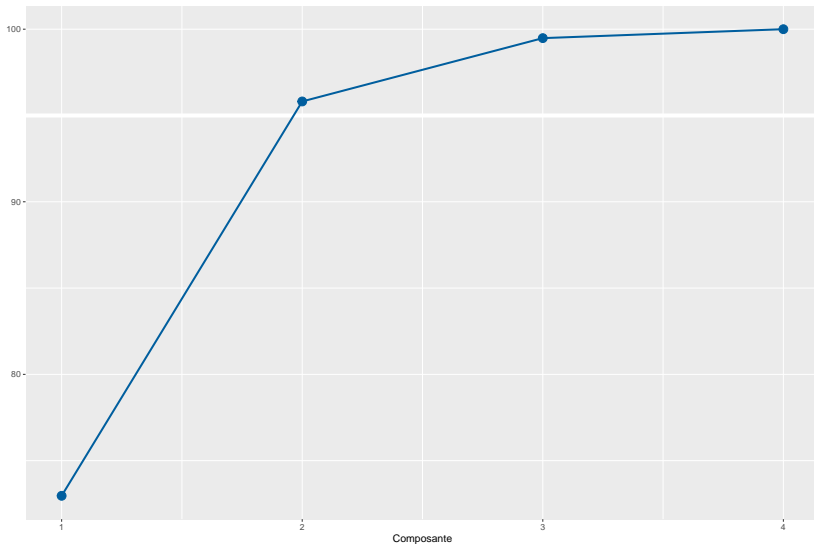
Proportion de variance expliquée



■ Choix du nombre d'axes



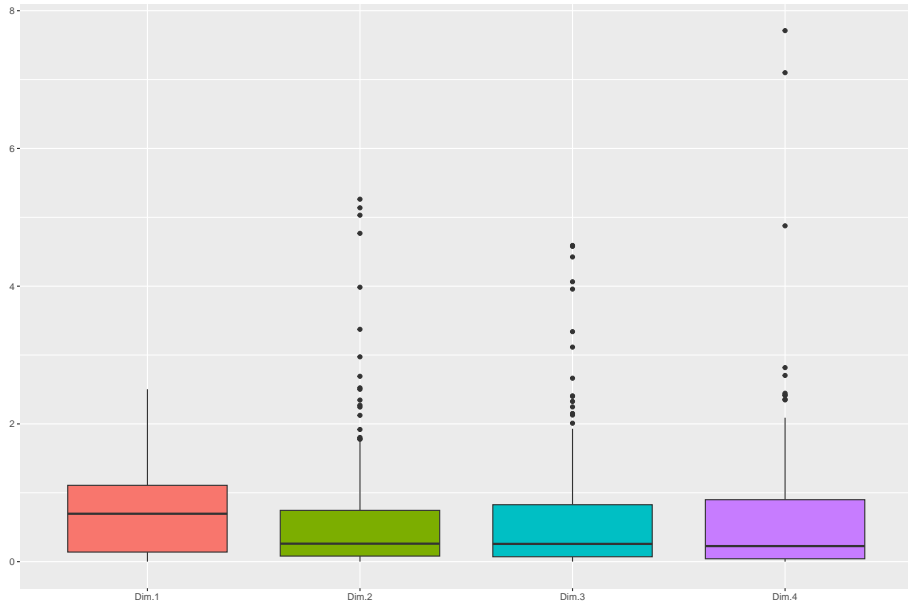
Proportion cumulée de variance expliquée



Notations X matrice $n \times p$ de données (déterministe)
 $\Sigma = X^T X$ matrice $p \times p$ de covariance empirique
 $\lambda_1 > \lambda_2 > \dots > \lambda_p$ valeurs propres ordonnées de Σ
 $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_p$ vecteurs propres correspondants

$$\frac{\langle \vec{x}^i, \vec{w}_k \rangle^2}{n\lambda_k} \quad \text{contribution de l'individu } i \text{ à l'axe } k$$

■ Contribution des individus



Notations X matrice $n \times p$ de données (déterministe)
 $\Sigma = X^T X$ matrice $p \times p$ de covariance empirique
 $\lambda_1 > \lambda_2 > \dots > \lambda_p$ valeurs propres ordonnées de Σ
 $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_p$ vecteurs propres correspondants

$$\frac{\langle \vec{x}_j, X \vec{w}_k \rangle^2}{\sqrt{\lambda_k}} \quad \text{corrélation entre la variable } j \text{ et la composante } k$$

■ Cercle des corrélations

