Statistiques pour données de comptage

Bastien Batardière

INRAe

January 21, 2025

Données de comptage

$$\boldsymbol{Y} = \begin{array}{c} \xrightarrow{\quad\text{variables}\quad} \\ \begin{pmatrix} 12 & 0 & \cdots & 0 & 9 \\ 2 & 0 & \cdots & 0 & 0 \\ \vdots & & & & \vdots \\ 341 & 5 & \cdots & 1 & 0 \end{pmatrix} \end{array} \Bigg\downarrow \text{individus}$$

Adaptation des tâches statistiques: PLN

$$\left.\begin{array}{l} \\ \\ \\ \end{array}\right\} Y_i \sim \mathcal{P}(\exp(Z_i))$$

Adaptation des tâches statistiques: PLN

Covariance

$$\left.\begin{array}{l} Z_i \sim \mathcal{N}(\mu, \Sigma) \\ Y_i \sim \mathcal{P}(\exp(Z_i)) \end{array}\right\}$$

Adaptation des tâches statistiques: PLN

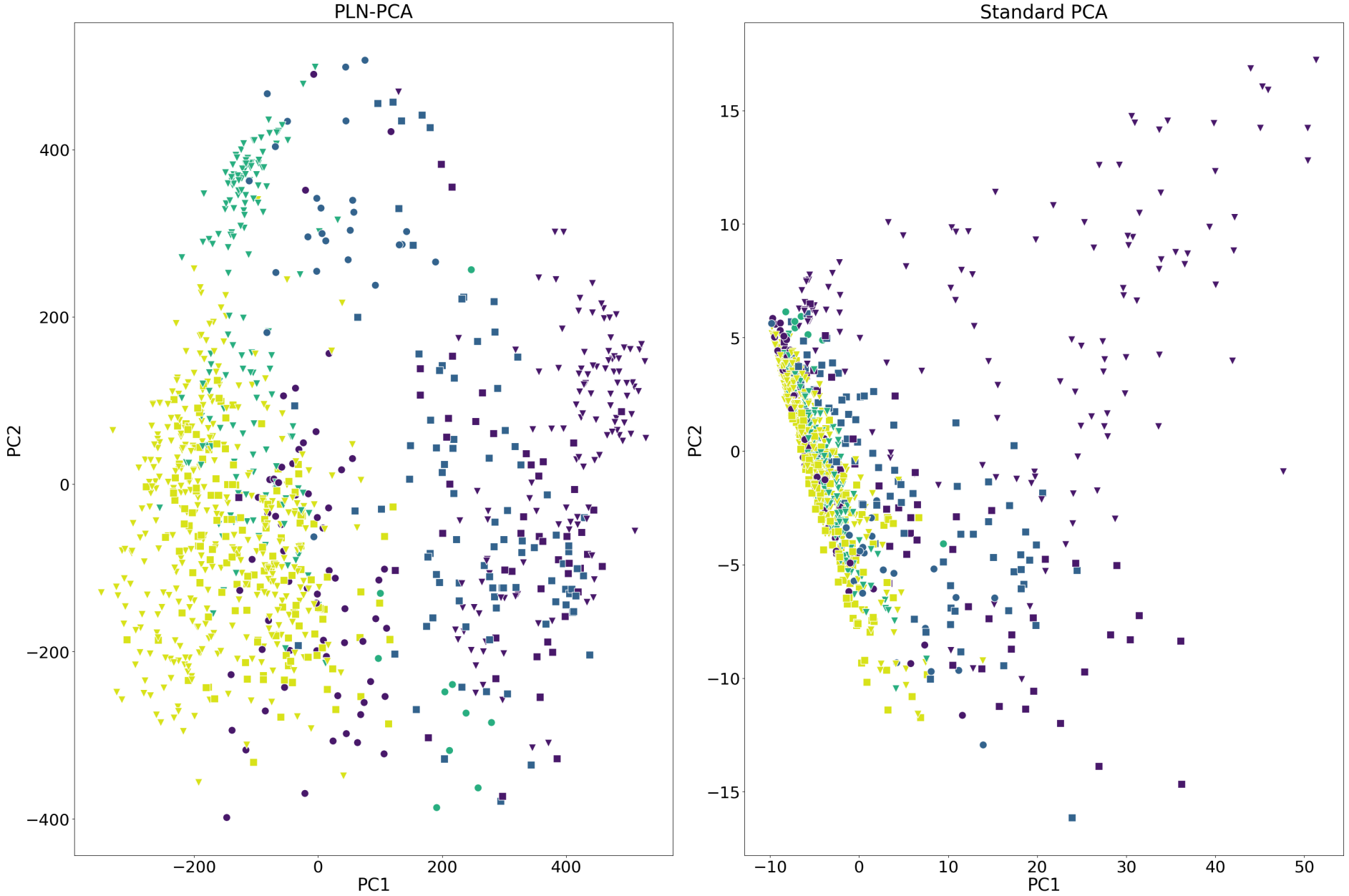
Covariance $\quad Z_i \sim \mathcal{N}(\mu, \Sigma)$

Séries temporelles $\quad Z_i \sim \mathcal{N}(AZ_{i-1}, \Sigma)$

$$\Big\} \; Y_i \sim \mathcal{P}(\exp(Z_i))$$

Adaptation des tâches statistiques: PLN

$$
\left.
\begin{array}{ll}
\text{Covariance} & Z_i \sim \mathcal{N}(\mu, \Sigma) \\
\text{Séries temporelles} & Z_i \sim \mathcal{N}(AZ_{i-1}, \Sigma) \\
\text{Clustering} & Z_i \sim \sum_{k=1}^{K} \alpha_k \mathcal{N}(\mu_k, \Sigma_k)
\end{array}
\right\} Y_i \sim \mathcal{P}(\exp(Z_i))
$$


PLN-PCA
Standard PCA
PC1
PC2
NK_cells
Monocytes
T_cells_CD4+
T_cells_CD8+
B_and_plasma_cells
Dendritic_cells
Mast_cells
Macrophages
Other_myeloid

```
pip install pyPLNmodels
```