Distributions sur le simplexe

Mahendra Mariadassou

INRAE - MaIAGE

JES 2022 - Fréjus

Définitions et résultats techniques

Moyenne et variance dans le simplexe

La moyenne et la variance sont définies à l'aide de la **philosophie de la transformation**

Definition

Soit $\mathbf{X}$ une composition aléatoire. La **moyenne** de $\mathbf{X}$ dans le simplexe, aussi appelée *centre de la composition*, et sa matrice de **(clr)-covariance** sont données par:

$$\mathbb{E}^{\oplus}[\mathbf{X}] = \mathsf{clr}^{-1}\left(\mathbb{E}[\mathsf{clr}(\mathbf{X})]\right) \quad \text{et} \quad \mathbb{V}^{\oplus}[\mathbf{X}] = \mathbb{V}[\mathsf{clr}(\mathbf{X})] = \mathbb{V}[\mathbf{G}_D \log(\mathbf{X})].$$

Moyenne et variance dans le simplexe

La moyenne et la variance sont définies à l'aide de la **philosophie de la transformation**

Definition

Soit $\mathbf{X}$ une composition aléatoire. La **moyenne** de $\mathbf{X}$ dans le simplexe, aussi appelée *centre de la composition*, et sa matrice de **(clr)-covariance** sont données par:

$$\mathbb{E}^{\oplus}[\mathbf{X}] = \mathsf{clr}^{-1}\left(\mathbb{E}[\mathsf{clr}(\mathbf{X})]\right) \quad \text{et} \quad \mathbb{V}^{\oplus}[\mathbf{X}] = \mathbb{V}[\mathsf{clr}(\mathbf{X})] = \mathbb{V}[\mathbf{G}_D \log(\mathbf{X})].$$

Note

- $\mathbb{E}^{\oplus}[\mathbf{X}]$ est à valeurs dans $\mathcal{S}^D$

Moyenne et variance dans le simplexe

La moyenne et la variance sont définies à l'aide de la **philosophie de la transformation**

Definition

Soit $\mathbf{X}$ une composition aléatoire. La **moyenne** de $\mathbf{X}$ dans le simplexe, aussi appelée *centre de la composition*, et sa matrice de **(clr)-covariance** sont données par:

$$\mathbb{E}^{\oplus}[\mathbf{X}] = \mathsf{clr}^{-1}\left(\mathbb{E}[\mathsf{clr}(\mathbf{X})]\right) \quad \text{et} \quad \mathbb{V}^{\oplus}[\mathbf{X}] = \mathbb{V}[\mathsf{clr}(\mathbf{X})] = \mathbb{V}[\mathbf{G}_D \log(\mathbf{X})].$$

Note

- $\mathbb{E}^{\oplus}[\mathbf{X}]$ est à valeurs dans $\mathcal{S}^D$
- $\mathbb{V}^{\oplus}[\mathbf{X}]$ est à valeurs dans $\mathcal{A}_D$, l'espace des matrices doublement centrées.

Inversion dans $\mathcal{A}_D$

Definition

Soit Σ une matrice symétrique de taille D. Σ est dite sous *forme canonique* sur le simplexe si elle est centrée en ligne et en colonne, c'est à dire si $\Sigma = \mathbf{G}_D \Sigma \mathbf{G}_D$.

Inversion dans $\mathcal{A}_D$

Definition

Soit Σ une matrice symétrique de taille D. Σ est dite sous *forme canonique* sur le simplexe si elle est centrée en ligne et en colonne, c'est à dire si $\Sigma = \mathbf{G}_D \Sigma \mathbf{G}_D$.

> $\mathbf{G}_D$ est de rang $D-1$ donc les matrices de clr-covariance ont un rang $\leq D-1$.
> Il faut donc redéfinir la notion d'inverse pour ces matrices.

Inversion dans $\mathcal{A}_D$

Definition

Soit Σ une matrice symétrique de taille D. Σ est dite sous *forme canonique* sur le simplexe si elle est centrée en ligne et en colonne, c'est à dire si $\Sigma = \mathbf{G}_D \Sigma \mathbf{G}_D$.

$\mathbf{G}_D$ est de rang $D-1$ donc les matrices de clr-covariance ont un rang $\leq D-1$.
Il faut donc redéfinir la notion d'inverse pour ces matrices.

Definition

Soit Σ une matrice symétrique de taille D sous forme canonique. Σ est dite **inversible** sur le simplexe s'il existe une base $\mathbf{V}$ telle que $\Sigma^{\star} = \mathbf{V}^T \Sigma \mathbf{V}$ est inversible. Dans ce cas, l'inverse de Σ dans le simplexe est $\Sigma^{-1} = \mathbf{V} {\Sigma^{\star}}^{-1} \mathbf{V}^T$.

Inversion dans $\mathcal{A}_D$

Definition

Soit Σ une matrice symétrique de taille D. Σ est dite sous *forme canonique* sur le simplexe si elle est centrée en ligne et en colonne, c'est à dire si $\Sigma = \mathbf{G}_D \Sigma \mathbf{G}_D$.

$\mathbf{G}_D$ est de rang $D-1$ donc les matrices de clr-covariance ont un rang $\leq D-1$. Il faut donc redéfinir la notion d'inverse pour ces matrices.

Definition

Soit Σ une matrice symétrique de taille D sous forme canonique. Σ est dite **inversible** sur le simplexe s'il existe une base $\mathbf{V}$ telle que $\Sigma^\star = \mathbf{V}^T \Sigma \mathbf{V}$ est inversible. Dans ce cas, l'inverse de Σ dans le simplexe est $\Sigma^{-1} = \mathbf{V} {\Sigma^\star}^{-1} \mathbf{V}^T$.

On impose Σ sous forme canonique uniquement pour alléger les notations, sinon:

$$\mathbf{V}^T \Sigma \mathbf{V} = \mathbf{V}^T \mathbf{G}_D \Sigma \mathbf{G}_D \mathbf{V}.$$

Inversion dans $\mathscr{A}_D$ (II)

L'inversibilité ne dépend pas du choix de la base $\mathbf{V}$.

Inversion dans $\mathcal{A}_D$ (II)

L'inversibilité ne dépend pas du choix de la base $\mathbf{V}$.

Theorem

Soit Σ une matrice symétrique inversible dans le simplexe. L'inverse de Σ dans le simplexe ne dépend pas du choix de $\mathbf{V}$ et s'écrit:

$$\Sigma^{-1} = (\mathbf{G}_D \Sigma \mathbf{G}_D + \mathbf{1}_D \mathbf{1}_D^T)^{-1} - \mathbf{1}_D \mathbf{1}_D^T$$

Preuve

Par simplicité, on suppose Σ est sous forme canonique. Soit $\mathbf{V}$ une base et $\Sigma^\star = \mathbf{V}^T \Sigma \mathbf{V}$. Pour travailler uniquement avec des matrices carrés, on introduit la matrice orthogonale $\mathbf{P}_V = [\mathbf{V} \frac{1}{\sqrt{D}} \mathbf{1}_D]$ de sorte que

$$\Sigma = \mathbf{V} \Sigma^\star \mathbf{V}^T = \mathbf{P}_V \begin{bmatrix} \Sigma^\star & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \mathbf{P}_V^T \text{ et } \Sigma^{-1} = \mathbf{V} {\Sigma^\star}^{-1} \mathbf{V}^T = \mathbf{P}_V \begin{bmatrix} {\Sigma^\star}^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \mathbf{P}_V^T.$$

Preuve

Par simplicité, on suppose Σ est sous forme canonique. Soit $\mathbf{V}$ une base et $\Sigma^\star = \mathbf{V}^T \Sigma \mathbf{V}$. Pour travailler uniquement avec des matrices carrés, on introduit la matrice orthogonale $\mathbf{P}_V = [\mathbf{V} \frac{1}{\sqrt{D}} \mathbf{1}_D]$ de sorte que

$$\Sigma = \mathbf{V}\Sigma^\star\mathbf{V}^T = \mathbf{P}_V \begin{bmatrix} \Sigma^\star & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \mathbf{P}_V^T \text{ et } \Sigma^{-1} = \mathbf{V}{\Sigma^\star}^{-1}\mathbf{V}^T = \mathbf{P}_V \begin{bmatrix} {\Sigma^\star}^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \mathbf{P}_V^T.$$

On conclut en notant

$$\begin{aligned} \Sigma^{-1} = \mathbf{P}_V \begin{bmatrix} {\Sigma^\star}^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \mathbf{P}_V^T &= \mathbf{P}_V \begin{bmatrix} {\Sigma^\star}^{-1} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{P}_V^T - \mathbf{P}_V \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{P}_V^T \\ &= \left(\mathbf{P}_V \begin{bmatrix} \Sigma^\star & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{P}_V^T \right)^{-1} - \mathbf{1}_D \mathbf{1}_D^T \\ &= (\Sigma + \mathbf{1}_D\mathbf{1}_D^T)^{-1} - \mathbf{1}_D\mathbf{1}_D^T \end{aligned}$$

Distribution de Dirichlet

Définition

Intuition

La *loi de Dirichlet* est la distribution d'une composition obtenue en tirant D composantes indépendamment suivant des lois Gamma partageant le même paramètre d'échelle avant de leur appliquer l'opérateur de clôture.

Définition

Intuition

La *loi de Dirichlet* est la distribution d'une composition obtenue en tirant D composantes indépendamment suivant des lois Gamma partageant le même paramètre d'échelle avant de leur appliquer l'opérateur de clôture.

Definition

Une composition aléatoire $\mathbf{X}$ suit une loi de Dirichlet $\mathcal{D}(\alpha)$ de paramètre $\alpha = (\alpha_1, \ldots, \alpha_D)$, s'il existe D variables indépendantes $Y_1, \ldots, Y_D$ de loi $Y_j \sim \Gamma(\alpha_j, 1)$ telles que

$$\mathbf{X} = \mathcal{C}(Y_1, \ldots, Y_D).$$

La densité de $\mathbf{X}$ par rapport à la mesure de Lebesgue est donnée par

$$f(\mathbf{x}; \alpha) = \frac{\Gamma(\alpha_0)}{\Pi_{j=1}^{D} \Gamma(\alpha_j)} \Pi_{j=1}^{D} x_i^{\alpha_i - 1} 1_{\{\mathbf{x} \in \mathcal{S}^D\}}$$

où $\Gamma(x) = \int_0^\infty t^{x-1} e^{-x} dx$ désigne la fonction gamma et $\alpha_0 = \alpha_1 + \cdots + \alpha_D$.

Remarques et propriétés

Il s'agit d'une généralisation multivariée de la loi Beta sur $[0, 1]$.

Remarques et propriétés

Il s'agit d'une généralisation multivariée de la loi Beta sur $[0, 1]$.

Si $\alpha_1 = \cdots = \alpha_D = \alpha$, on parle de loi de Dirichlet **symétrique**.

Remarques et propriétés

Il s'agit d'une généralisation multivariée de la loi Beta sur $[0, 1]$.

Si $\alpha_1 = \cdots = \alpha_D = \alpha$, on parle de loi de Dirichlet **symétrique**.

La loi de Dirichlet est **compatible** avec l'amalgamation. Si $(X_1, \dots, X_D) \sim \mathcal{D}(\alpha_1, \dots, \alpha_D)$, alors $\mathbf{X}' = (X_1 + X_2, X_3, \dots, X_D) \sim \mathcal{D}(\alpha_1 + \alpha_2, \alpha_3, \dots, \alpha_D)$.

Remarques et propriétés

Il s'agit d'une généralisation multivariée de la loi Beta sur $[0, 1]$.

Si $\alpha_1 = \cdots = \alpha_D = \alpha$, on parle de loi de Dirichlet **symétrique**.

La loi de Dirichlet est **compatible** avec l'amalgamation. Si $(X_1, \ldots, X_D) \sim \mathcal{D}(\alpha_1, \ldots, \alpha_D)$, alors $\mathbf{X}' = (X_1 + X_2, X_3, \ldots, X_D) \sim \mathcal{D}(\alpha_1 + \alpha_2, \alpha_3, \ldots, \alpha_D)$.

- Lorsque $\alpha_1 = \cdots = \alpha_D = 1$, $\mathcal{D}(\alpha)$ est la **loi uniforme** sur $\mathcal{S}^D$.
- Lorsque $\alpha_1 = \cdots = \alpha_D = 1/2$, $\mathcal{D}(\alpha)$ est la **loi de Jeffrey** sur le simplexe.

Moyenne et variance

Les moments de $\mathbf{X}$ ont des formes plus ou moins simples.

Moyenne et variance

Les moments de $\mathbf{X}$ ont des formes plus ou moins simples.

Moments arithmétiques

$$\mathbb{E}[\mathbf{X}] = \mathcal{C}(\alpha) = \bar{\alpha}$$
$$\mathbb{V}[\mathbf{X}] = \frac{1}{\alpha_0 + 1}\left(\mathsf{diag}(\bar{\alpha}) - \bar{\alpha}\bar{\alpha}^T\right) = \frac{1}{\alpha_0 + 1}\mathbf{G}_D \,\mathrm{diag}(\bar{\alpha})\mathbf{G}_D$$

où diag$(\mathbf{x})$ désigne la matrice diagonale de diagonale $\mathbf{x}$.

Moyenne et variance

Les moments de $\mathbf{X}$ ont des formes plus ou moins simples.

Moments arithmétiques

$$\mathbb{E}[\mathbf{X}] = \mathcal{C}(\alpha) = \bar{\alpha}$$
$$\mathbb{V}[\mathbf{X}] = \frac{1}{\alpha_0 + 1}\left(\mathrm{diag}(\bar{\alpha}) - \bar{\alpha}\bar{\alpha}^T\right) = \frac{1}{\alpha_0 + 1}\mathbf{G}_D\,\mathrm{diag}(\bar{\alpha})\mathbf{G}_D$$

où $\mathrm{diag}(\mathbf{x})$ désigne la matrice diagonale de diagonale $\mathbf{x}$.

Moments compositionnels (Aitchison 1986)

$$\mathbb{E}^{\oplus}[\mathbf{X}] = \mathrm{clr}^{-1}(\psi(\alpha))$$
$$\mathbb{V}^{\oplus}[\mathbf{X}] = \mathbf{G}_D\mathrm{diag}(\psi'(\alpha))\mathbf{G}_D$$

où $\psi(x) = \Gamma'(x)/\Gamma(x)$ (resp. $\psi'(x)$) est la fonction digamma (resp. trigamma)

Moyenne et variance

Les moments de $\mathbf{X}$ ont des formes plus ou moins simples.

Moments arithmétiques

$$\mathbb{E}[\mathbf{X}] = \mathcal{C}(\alpha) = \bar{\alpha}$$
$$\mathbb{V}[\mathbf{X}] = \frac{1}{\alpha_0 + 1}\left(\text{diag}(\bar{\alpha}) - \bar{\alpha}\bar{\alpha}^T\right) = \frac{1}{\alpha_0 + 1}\mathbf{G}_D \,\mathrm{diag}(\bar{\alpha})\mathbf{G}_D$$

où $\text{diag}(\mathbf{x})$ désigne la matrice diagonale de diagonale $\mathbf{x}$.

Moments compositionnels (Aitchison 1986)

$$\mathbb{E}^{\oplus}[\mathbf{X}] = \text{clr}^{-1}(\psi(\alpha))$$
$$\mathbb{V}^{\oplus}[\mathbf{X}] = \mathbf{G}_D \text{diag}(\psi'(\alpha))\mathbf{G}_D$$

où $\psi(x) = \Gamma'(x)/\Gamma(x)$ (resp. $\psi'(x)$) est la fonction digamma (resp. trigamma)

> Les deux variances $\mathbb{V}[\mathbf{X}]$ et $\mathbb{V}^{\oplus}[\mathbf{X}]$ sont *double-centrées*.

Remarques (bis)

- Au centrage près, les deux variances sont **diagonales**.
- Pas de possibilité de modéliser la dépendance entre composantes.

Remarques (bis)

- Au centrage près, les deux variances sont **diagonales**.
- Pas de possibilité de modéliser la dépendance entre composantes.

La décomposition $\alpha = \alpha_0 \bar{\alpha}$ fait apparaître

- une composition $\bar{\alpha}$ ($\simeq$ composition centrale)
- un facteur d'échelle α_0 (paramètre de dispersion)

Plus α_0 est grand, plus les échantillons sont proches de $\bar{\alpha}$

$\mathcal{D}(\alpha_0\bar{\alpha})$ avec $\bar{\alpha} = (0.5, 0.3, 0.2)$ et 2 valeurs de α_0

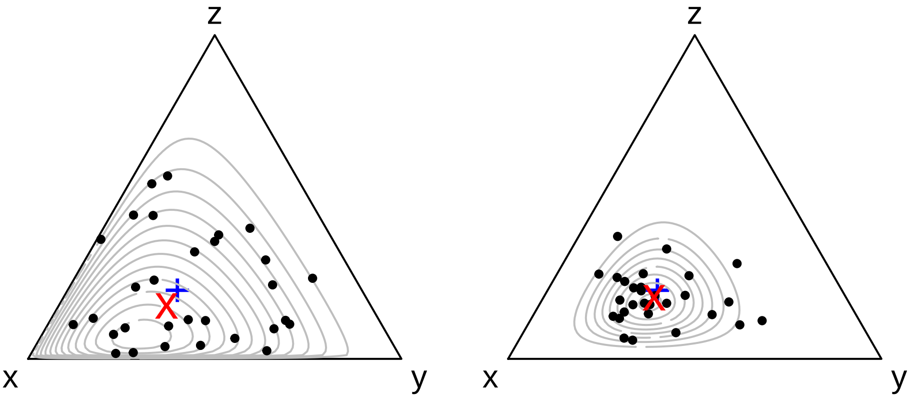

Figure 1: $\alpha_0 = 6$ (gauche) et $\alpha_0 = 18$ (droite)

Moyennes arithmétique ($+$), géométrique ($\times$) et isocontours (-)

Loi normale sur le simplexe

Définition (Figueras 2003)

La **loi normale sur le simplexe** est définie à l'aide de la philosophie de la transformation.

Définition (Figueras 2003)

La **loi normale sur le simplexe** est définie à l'aide de la philosophie de la transformation.

Definition

Soit $\mathbf{X}$ une composition aléatoire à valeurs dans $\mathcal{S}^D$. On dit que $\mathbf{X}$ suit une loi normale sur le simplexe si et seulement si, pour tout matrice de base $\mathbf{V}$ le vecteur $\mathbf{X}^\star = \text{ilr}_{\mathbf{V}}(\mathbf{X})$ suit une loi normale multivariée sur $\mathbb{R}^{D-1}$.

Définition (Figueras 2003)

La **loi normale sur le simplexe** est définie à l'aide de la philosophie de la transformation.

Definition

Soit $\mathbf{X}$ une composition aléatoire à valeurs dans $\mathcal{S}^D$. On dit que $\mathbf{X}$ suit une loi normale sur le simplexe si et seulement si, pour tout matrice de base $\mathbf{V}$ le vecteur $\mathbf{X}^\star = \text{ilr}_\mathbf{V}(\mathbf{X})$ suit une loi normale multivariée sur $\mathbb{R}^{D-1}$.

- *Quid* des paramètres de la distribution ?
- Dépendent-ils du choix de $\mathbf{V}$?

Remarques

Supposons que $\text{ilr}_{\mathbf{V}}(\mathbf{X}) \sim \mathcal{N}(\mu_V^\star, \Sigma_V^\star)$, on a

$$\text{clr}(\mathbf{x}) = \mathbf{V}\,\text{ilr}_V(\mathbf{x}) \sim \mathcal{N}(\underbrace{\mathbf{V}\mu_V^\star}_{\mu}, \underbrace{\mathbf{V}\Sigma_V^\star\mathbf{V}^T}_{\Sigma})$$

Remarques

Supposons que $\mathrm{ilr}_{\mathbf{V}}(\mathbf{X}) \sim \mathcal{N}(\mu_V^\star, \Sigma_V^\star)$, on a

$$\mathrm{clr}(\mathbf{x}) = \mathbf{V}\,\mathrm{ilr}_V(\mathbf{x}) \sim \mathcal{N}(\underbrace{\mathbf{V}\mu_V^\star}_{\mu}, \underbrace{\mathbf{V}\Sigma_V^\star\mathbf{V}^T}_{\Sigma})$$

μ et Σ sont des **invariants** de $\mathbf{X}$. En particulier

$$\begin{aligned} \mu &= \mathrm{clr}(\mathbb{E}^\oplus[\mathbf{X}]) \\ \Sigma &= \mathbb{V}^\oplus[\mathbf{X}] \end{aligned}$$

Définition (II) (Van den Boogaart and Tolosana-Delgado 2013)

Une composition aléatoire $\mathbf{X}$ a une *distribution normale sur le simplexe* de moyenne μ et variance Σ, notée $\mathcal{N}_{\mathcal{S}}(\mu, \Sigma)$, si son produit scalaire $\langle \mathbf{X}, \mathbf{u} \rangle_A$ avec *chaque* composition du simplexe suit une distribution normale de moyenne $\mu^T \operatorname{clr}(\mathbf{u})$ et de variance clr$(\mathbf{u})\Sigma$clr$(\mathbf{u})^T$.

Définition (II) (Van den Boogaart and Tolosana-Delgado 2013)

Une composition aléatoire $\mathbf{X}$ a une *distribution normale sur le simplexe* de moyenne μ et variance Σ, notée $\mathcal{N}_{\mathcal{S}}(\mu, \Sigma)$, si son produit scalaire $\langle \mathbf{X}, \mathbf{u} \rangle_A$ avec *chaque* composition du simplexe suit une distribution normale de moyenne $\mu^T \operatorname{clr}(\mathbf{u})$ et de variance $\text{clr}(\mathbf{u})\Sigma\text{clr}(\mathbf{u})^T$.

En particulier, si on considère une base $\mathbf{V}$, le vecteur $\mathbf{X}^\star = \text{ilr}_{\mathbf{V}}(\mathbf{X})$ suit une loi $\mathcal{N}(\mu^\star, \Sigma^\star)$ avec $\mu^\star = \text{ilr}_{\mathbf{V}}(\mu)$ et $\Sigma^\star = \mathbf{V}^T \Sigma \mathbf{V}$.

Définition (II) (Van den Boogaart and Tolosana-Delgado 2013)

Une composition aléatoire $\mathbf{X}$ a une *distribution normale sur le simplexe* de moyenne μ et variance Σ, notée $\mathcal{N}_{\mathcal{S}}(\mu, \Sigma)$, si son produit scalaire $\langle \mathbf{X}, \mathbf{u} \rangle_A$ avec *chaque* composition du simplexe suit une distribution normale de moyenne $\mu^T \operatorname{clr}(\mathbf{u})$ et de variance $\text{clr}(\mathbf{u})\Sigma\text{clr}(\mathbf{u})^T$.

En particulier, si on considère une base $\mathbf{V}$, le vecteur $\mathbf{X}^\star = \text{ilr}_{\mathbf{V}}(\mathbf{X})$ suit une loi $\mathcal{N}(\mu^\star, \Sigma^\star)$ avec $\mu^\star = \text{ilr}_{\mathbf{V}}(\mu)$ et $\Sigma^\star = \mathbf{V}^T \Sigma \mathbf{V}$.

La densité de $\mathbf{X}$ par rapport à la mesure-image de la mesure de Lebesgue sur $\mathbb{R}^{D-1}$ (aussi appelée mesure de Aitchinson) est:

$$f(\mathbf{x}; \mu^\star, \Sigma^\star) = \frac{1}{\sqrt{(2\pi)^{D-1} |\Sigma^\star|}} \exp\left[-\frac{1}{2} (\text{ilr}_{\mathbf{V}}(\mathbf{x}) - \mu^\star)^T {\Sigma^\star}^{-1} (\text{ilr}_{\mathbf{V}}(\mathbf{x}) - \mu^\star) \right].$$

Remarques (II)

Le passage par $(\mu^\star, \Sigma^\star)$ facilite l'écriture en garantissant que $|\Sigma^\star|$ est non-nul mais on pourrait s'en passer.

Remarques (II)

Le passage par $(\mu^\star, \Sigma^\star)$ facilite l'écriture en garantissant que $|\Sigma^\star|$ est non-nul mais on pourrait s'en passer.

Sans contraintes supplémentaires, (μ, Σ) ne sont pas identifiables. Ils le deviennent si impose $\mu^T \mathbf{1}_D = 0$ et $\mathbf{\Sigma} \mathbf{1}_D = \mathbf{0}_D$.

Remarques (II)

Le passage par $(\mu^\star, \Sigma^\star)$ facilite l'écriture en garantissant que $|\Sigma^\star|$ est non-nul mais on pourrait s'en passer.

Sans contraintes supplémentaires, (μ, Σ) ne sont pas identifiables. Ils le deviennent si impose $\mu^T \mathbf{1}_D = 0$ et $\Sigma \mathbf{1}_D = \mathbf{0}_D$.

Sans contraintes supplémentaires sur (μ, Σ) on a

$$\mathrm{clr}(\mathbb{E}^\oplus[\mathbf{X}]) = \mathbf{G}_D \mu \quad \text{et} \quad \mathbb{V}^\oplus[\mathbf{X}] = \mathbf{G}_D \Sigma \mathbf{G}_D$$

Simulations

On peut facilement simuler des compositions aléatoires de loi normale sur le simplexe:

$$
\begin{aligned}
\mathbf{Z} &\sim \mathcal{N}(\mu^\star, \Sigma^\star) \\
\mathbf{X} &= \mathrm{ilr}_{\mathbf{V}}^{-1}(\mathbf{Z}).
\end{aligned}
$$

en se donnant une base $\mathbf{V}$ quelconque et en calculant $\mu^\star$ et $\Sigma^\star$.

Simulations

On peut facilement simuler des compositions aléatoires de loi normale sur le simplexe:

$$\mathbf{Z} \sim \mathcal{N}(\mu^\star, \Sigma^\star)$$
$$\mathbf{X} = \mathsf{ilr}_{\mathbf{V}}^{-1}(\mathbf{Z}).$$

en se donnant une base $\mathbf{V}$ quelconque et en calculant $\mu^\star$ et $\Sigma^\star$.

$$\mu^\star = (0, 0)$$
$$\Sigma^\star = \begin{bmatrix} 2 & \rho \\ \rho & 2 \end{bmatrix}$$
$$\mathbf{V} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$$

$\mathbf{V}$ est la base pivot usuelle et on considère $\rho \in \{-1.5, 0, 1.5\}$.

Illustrations ($\rho = -1.5$)

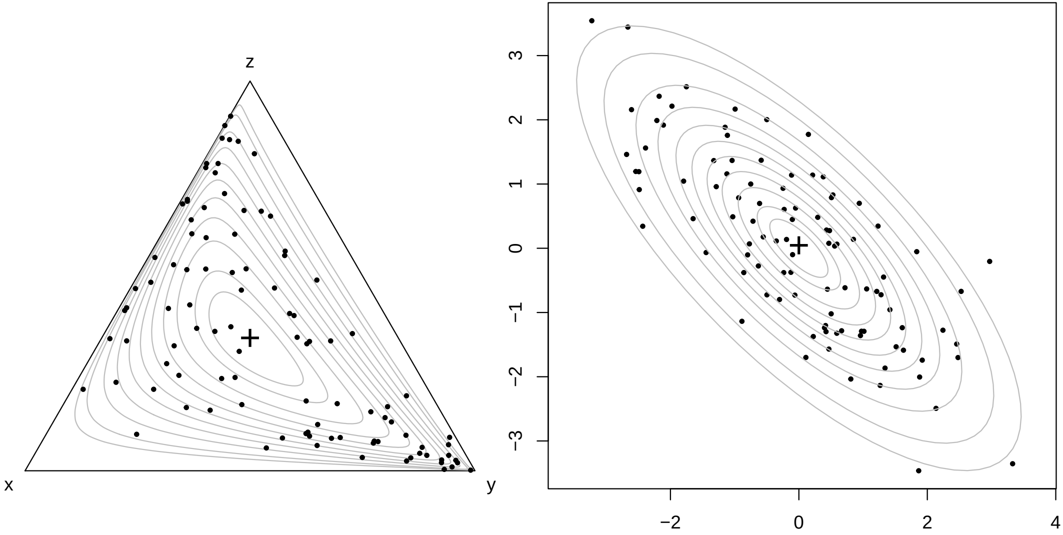

Figure 2: Moyenne de la distribution (+) et isocontours (-)

Illustrations ($\rho = 0$)

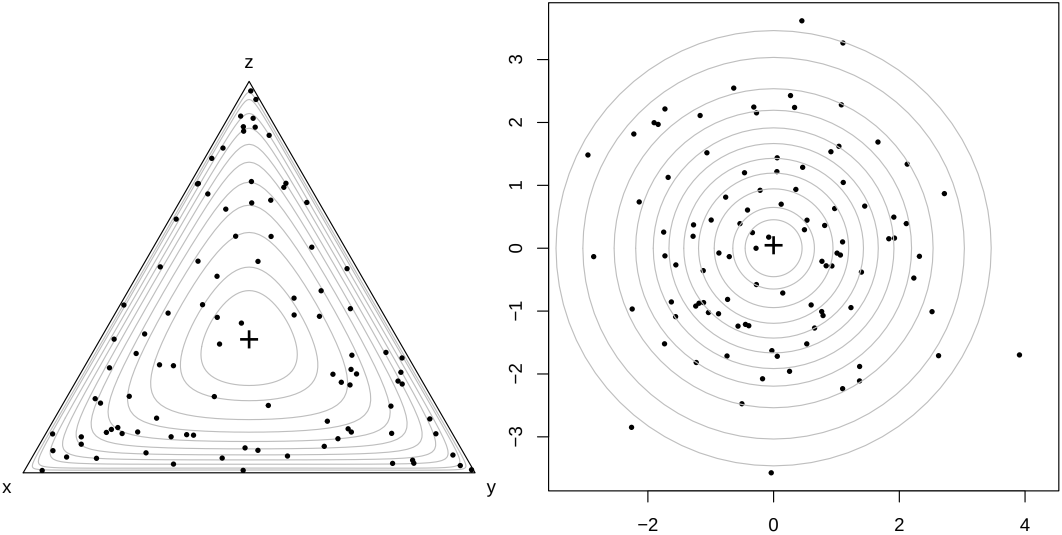

Figure 3: Moyenne de la distribution (+) et isocontours (-)

Illustrations ($\rho = 1.5$)

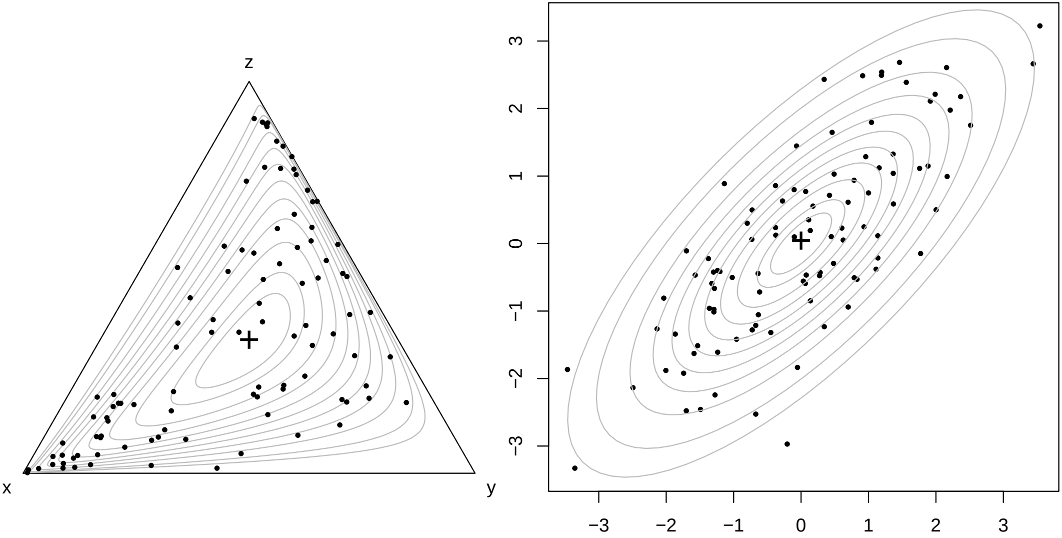

Figure 4: Moyenne de la distribution (+) et isocontours (-)

Illustrations $\mathbb{E}^{\oplus}[\mathbf{X}] = (1/6, 1/3, 1/2)$ et $\rho = -1.5$

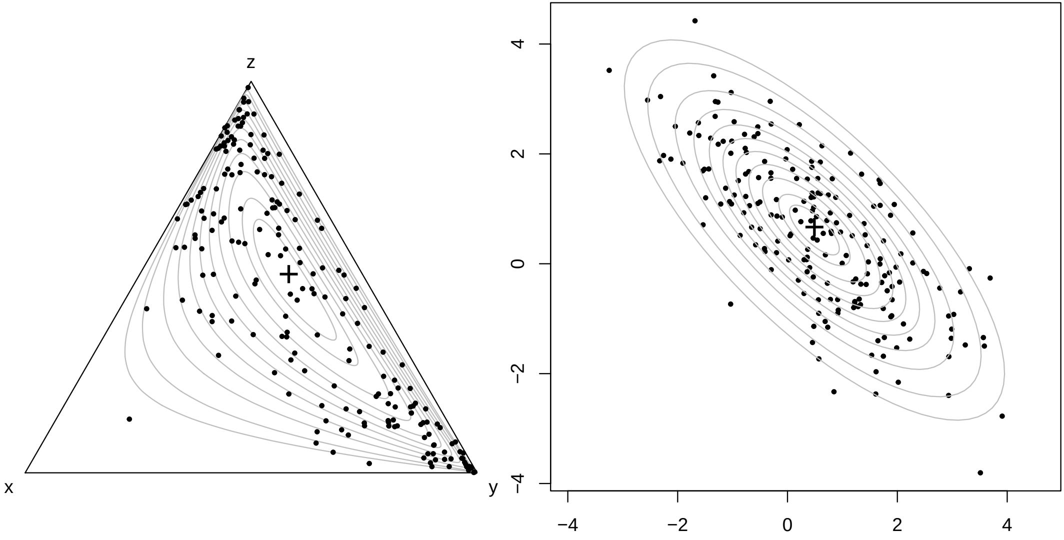

Figure 5: Moyenne de la distribution (+) et isocontours (-)

Illustrations $\mathbb{E}^{\oplus}[\mathbf{X}] = (1/6, 1/3, 1/2)$ et $\rho = 0$

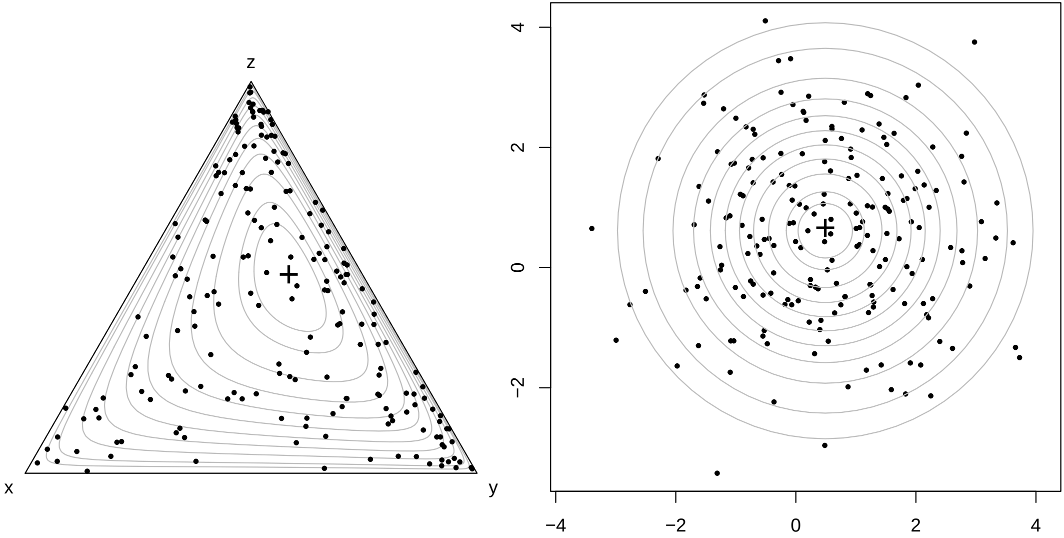

Figure 6: Moyenne de la distribution (+) et isocontours (-)

Illustrations $\mathbb{E}^{\oplus}[\mathbf{X}] = (1/6, 1/3, 1/2)$ et $\rho = 1.5$

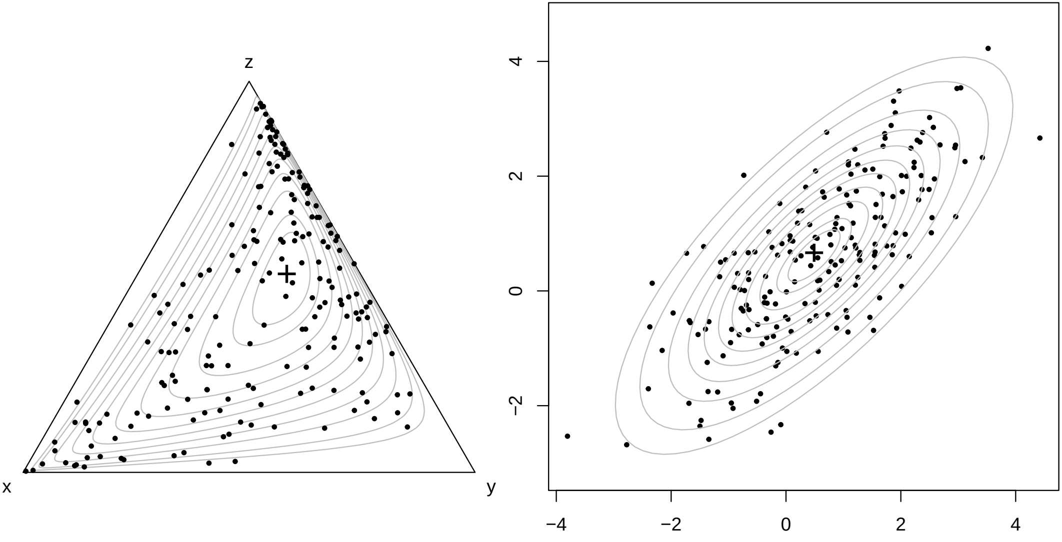

Figure 7: Moyenne de la distribution (+) et isocontours (-)

Loi elliptique sur le simplexe

Définition

La densité normale sur le simplexe peut être généralisée aux familles elliptiques pour avoir des queues de distributions plus lourdes que la gaussienne (utile pour la détection d'atypiques).

Définition

La densité normale sur le simplexe peut être généralisée aux familles elliptiques pour avoir des queues de distributions plus lourdes que la gaussienne (utile pour la détection d'atypiques).

Definition

Une composition aléatoire $\mathbf{X}$ a une *distribution elliptique sur le simplexe* de moyenne μ et variance Σ si n'importe laquelle de ses transformées ilr suit une distribution elliptique sur $\mathbb{R}^{D-1}$.

Définition

La densité normale sur le simplexe peut être généralisée aux familles elliptiques pour avoir des queues de distributions plus lourdes que la gaussienne (utile pour la détection d'atypiques).

Definition

Une composition aléatoire $\mathbf{X}$ a une *distribution elliptique sur le simplexe* de moyenne μ et variance Σ si n'importe laquelle de ses transformées ilr suit une distribution elliptique sur $\mathbb{R}^{D-1}$.

Dans la base $\mathbf{V}$ avec $\mu^\star = \mathbf{V}^T\mu$ et $\Sigma^\star = \mathbf{V}^T\Sigma\mathbf{V}$, la densité de $\mathbf{X}^\star = \text{ilr}_{\mathbf{V}}(\mathbf{X})$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^{D-1}$ est:

$$f(\mathbf{x}^\star; \mu^\star, \Sigma^\star) = k(\mu^\star, \Sigma^\star) \times g\left((\mathbf{x}^\star - \mu^\star)^T {\Sigma^\star}^{-1} (\mathbf{x}^\star - \mu^\star)\right).$$

où $k(\mu^\star, \Sigma^\star)$ est la constante de normalisation et g une fonction à valeurs dans $\mathbb{R}_+$.

Remarques (I)

- Pour la loi gaussienne, $g(x) = \exp(-x^2/2)$
- Pour la loi de Student à ν degrés de libertés, $g(x) = (1 + x/\nu)^{-(\nu+D-1)/2}$

Remarques (I)

- Pour la loi gaussienne, $g(x) = \exp(-x^2/2)$
- Pour la loi de Student à ν degrés de libertés, $g(x) = (1 + x/\nu)^{-(\nu+D-1)/2}$

- La distribution est invariante au choix de $\mathbf{V}$
- Les paramètres μ et Σ ne sont identifiables que via $\mathbf{G}_D\mu)$ et $\mathbf{G}_D\Sigma\mathbf{G}_D$

Remarques (I)

- Pour la loi gaussienne, $g(x) = \exp(-x^2/2)$
- Pour la loi de Student à ν degrés de libertés, $g(x) = (1 + x/\nu)^{-(\nu+D-1)/2}$

- La distribution est invariante au choix de $\mathbf{V}$
- Les paramètres μ et Σ ne sont identifiables que via $\mathbf{G}_D\mu$) et $\mathbf{G}_D\Sigma\mathbf{G}_D$

Simuler une loi elliptique sur le simplexe est **aussi difficile** que de simuler une loi elliptique sur $\mathbb{R}^{D-1}$. Ce n'est pas toujours possible...

Simulations

On peut simuler des compositions aléatoires de **loi de student** sur le simplexe:

$$
\begin{aligned}
\mathbf{Z} &\sim \mathcal{N}(\mathbf{0}_{D-1}, \Sigma^\star) \\
u &\sim \chi^2_\nu \\
\mathbf{Y} &= \sqrt{\nu/u}\,\mathbf{Z} + \mu^\star \\
\mathbf{X} &= \mathrm{ilr}_{\mathbf{V}}^{-1}(\mathbf{Z}).
\end{aligned}
$$

en se donnant une base $\mathbf{V}$ quelconque et en calculant $\mu^\star$ et $\Sigma^\star$.

Simulations

On peut simuler des compositions aléatoires de **loi de student** sur le simplexe:

$$\begin{aligned}
\mathbf{Z} &\sim \mathcal{N}(\mathbf{0}_{D-1}, \Sigma^\star) \\
u &\sim \chi^2_\nu \\
\mathbf{Y} &= \sqrt{\nu/u}\mathbf{Z} + \mu^\star \\
\mathbf{X} &= \mathrm{ilr}_{\mathbf{V}}^{-1}(\mathbf{Z}).
\end{aligned}$$

en se donnant une base $\mathbf{V}$ quelconque et en calculant $\mu^\star$ et $\Sigma^\star$.

$$\begin{aligned}
\mu^\star &= (0, 0) \\
\Sigma^\star &= \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.5 \end{bmatrix} \\
\mathbf{V} &= \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{bmatrix}
\end{aligned}$$

$\mathbf{V}$ est la base pivot usuelle.

Illustrations ($\nu = 3$)

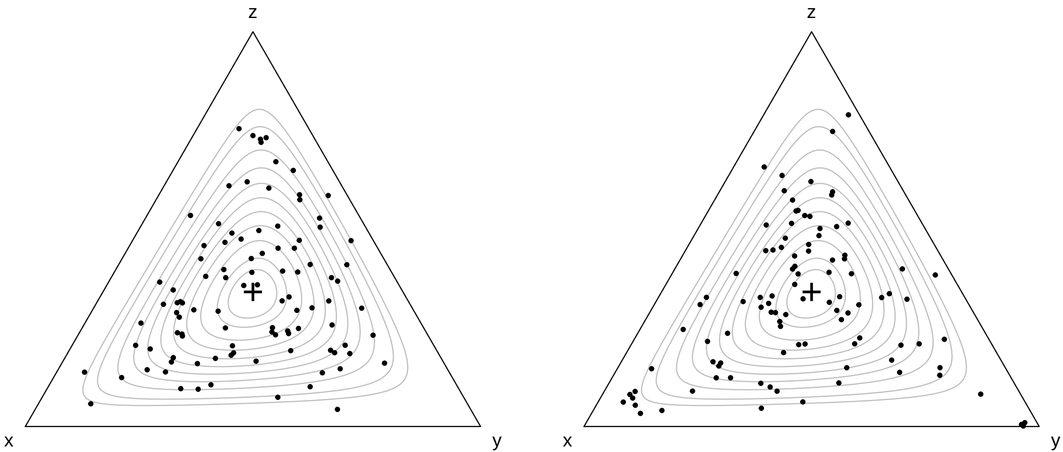

Figure 8: Loi normale (gauche) et de student (droite) de mêmes paramètres ($\nu = 3$ pour la student) sur le simplexe. Les isocontours sont celles de la gaussienne dans les deux cas.

Illustrations ($\nu = 20$)

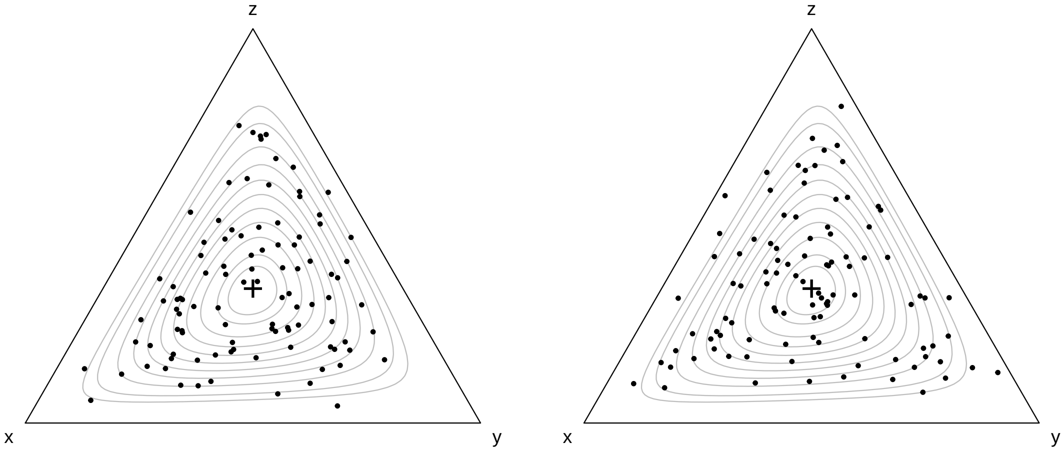

Figure 9: Loi normale (gauche) et de student (droite) de mêmes paramètres ($\nu = 20$ pour la student) sur le simplexe. Les isocontours sont celles de la gaussienne dans les deux cas.

Loi de Aitchison sur le simplexe

Définition (Aitchison 1985)

C'est une généralisation de la loi de Dirichlet et de la loi normale sur le simplexe.

Définition (Aitchison 1985)

C'est une généralisation de la loi de Dirichlet et de la loi normale sur le simplexe.

Definition

Soit Σ une matrice symétrique définie positive dans le simplexe, μ une composition et α un réel positif. La densité de Aitchison par rapport à la mesure de Aitchinson sur le simplexe $\mathcal{A}(\alpha, \mu, \Sigma)$ est donnée par

$$f(\mathbf{x}; \alpha, \mu, \Sigma) = e^{-\kappa(\alpha,\mu,\Sigma)} \times \Pi_{j=1}^{D} x_j^{\alpha-1} \times \exp\left[-\frac{1}{2}\mathsf{clr}(\mathbf{x} \ominus \mu)^T \Sigma^{-1} \mathsf{clr}(\mathbf{x} \ominus \mu)\right].$$

où $\Sigma^{-1} = (\mathbf{G}_D \Sigma \mathbf{G}_D + \mathbf{1}_D \mathbf{1}_D^T)^{-1} - \mathbf{1}_D \mathbf{1}_D^T$.

- μ est une *quasi-moyenne*

Définition (Aitchison 1985)

C'est une généralisation de la loi de Dirichlet et de la loi normale sur le simplexe.

Definition

Soit Σ une matrice symétrique définie positive dans le simplexe, μ une composition et α un réel positif. La densité de Aitchison par rapport à la mesure de Aitchinson sur le simplexe $\mathcal{A}(\alpha, \mu, \Sigma)$ est donnée par

$$f(\mathbf{x}; \alpha, \mu, \Sigma) = e^{-\kappa(\alpha,\mu,\Sigma)} \times \Pi_{j=1}^{D} x_j^{\alpha-1} \times \exp\left[-\frac{1}{2}\text{clr}(\mathbf{x} \ominus \mu)^T \Sigma^{-1} \text{clr}(\mathbf{x} \ominus \mu)\right].$$

où $\Sigma^{-1} = (\mathbf{G}_D \Sigma \mathbf{G}_D + \mathbf{1}_D \mathbf{1}_D^T)^{-1} - \mathbf{1}_D \mathbf{1}_D^T$.

- μ est une *quasi-moyenne*
- Σ une *quasi-variance*

Définition (Aitchison 1985)

C'est une généralisation de la loi de Dirichlet et de la loi normale sur le simplexe.

Definition

Soit Σ une matrice symétrique définie positive dans le simplexe, μ une composition et α un réel positif. La densité de Aitchison par rapport à la mesure de Aitchinson sur le simplexe $\mathcal{A}(\alpha, \mu, \Sigma)$ est donnée par

$$f(\mathbf{x}; \alpha, \mu, \Sigma) = e^{-\kappa(\alpha,\mu,\Sigma)} \times \Pi_{j=1}^{D} x_j^{\alpha-1} \times \exp\left[-\frac{1}{2}\mathsf{clr}(\mathbf{x} \ominus \mu)^T \Sigma^{-1} \mathsf{clr}(\mathbf{x} \ominus \mu)\right].$$

où $\Sigma^{-1} = (\mathbf{G}_D \Sigma \mathbf{G}_D + \mathbf{1}_D \mathbf{1}_D^T)^{-1} - \mathbf{1}_D \mathbf{1}_D^T$.

- μ est une *quasi-moyenne*
- Σ une *quasi-variance*
- $f(\mathbf{x}; \alpha, \mu, \Sigma)$ est le produit des densités d'une

Définition (Aitchison 1985)

C'est une généralisation de la loi de Dirichlet et de la loi normale sur le simplexe.

Definition

Soit Σ une matrice symétrique définie positive dans le simplexe, μ une composition et α un réel positif. La densité de Aitchison par rapport à la mesure de Aitchinson sur le simplexe $\mathcal{A}(\alpha,\mu,\Sigma)$ est donnée par

$$f(\mathbf{x};\alpha,\mu,\Sigma) = e^{-\kappa(\alpha,\mu,\Sigma)} \times \Pi_{j=1}^{D} x_j^{\alpha-1} \times \exp\left[-\frac{1}{2}\mathsf{clr}(\mathbf{x}\ominus\mu)^T\Sigma^{-1}\mathsf{clr}(\mathbf{x}\ominus\mu)\right].$$

où $\Sigma^{-1} = (\mathbf{G}_D\Sigma\mathbf{G}_D + \mathbf{1}_D\mathbf{1}_D^T)^{-1} - \mathbf{1}_D\mathbf{1}_D^T$.

- μ est une *quasi-moyenne*
- Σ une *quasi-variance*
- $f(\mathbf{x};\alpha,\mu,\Sigma)$ est le produit des densités d'une
  - d'une Dirichlet symétrique de paramètre α et

Définition (Aitchison 1985)

C'est une généralisation de la loi de Dirichlet et de la loi normale sur le simplexe.

Definition

Soit Σ une matrice symétrique définie positive dans le simplexe, μ une composition et α un réel positif. La densité de Aitchison par rapport à la mesure de Aitchinson sur le simplexe $\mathcal{A}(\alpha, \mu, \Sigma)$ est donnée par

$$f(\mathbf{x}; \alpha, \mu, \Sigma) = e^{-\kappa(\alpha,\mu,\Sigma)} \times \Pi_{j=1}^{D} x_j^{\alpha-1} \times \exp\left[-\frac{1}{2}\mathsf{clr}(\mathbf{x} \ominus \mu)^T \Sigma^{-1} \mathsf{clr}(\mathbf{x} \ominus \mu)\right].$$

où $\Sigma^{-1} = (\mathbf{G}_D \Sigma \mathbf{G}_D + \mathbf{1}_D \mathbf{1}_D^T)^{-1} - \mathbf{1}_D \mathbf{1}_D^T$.

- μ est une *quasi-moyenne*
- Σ une *quasi-variance*
- $f(\mathbf{x}; \alpha, \mu, \Sigma)$ est le produit des densités d'une
  - d'une Dirichlet symétrique de paramètre α et
  - d'une Normale sur le simplexe de paramètres (μ, Σ).

Remarques

- Quand $\alpha = 0$, $\mathcal{A}(\alpha, \mu, \Sigma) = \mathcal{N}_{\mathcal{S}}(\mu, \Sigma)$

Remarques

- Quand $\alpha = 0$, $\mathcal{A}(\alpha, \mu, \Sigma) = \mathcal{N}_{\mathcal{S}}(\mu, \Sigma)$
- Quand $\Sigma^{-1} \to \mathbf{0}_{D\times D}$, $\mathcal{A}(\alpha, \mu, \Sigma) \to \mathcal{D}(\alpha, \dots, \alpha)$ en distribution.

Remarques

- Quand $\alpha = 0$, $\mathcal{A}(\alpha, \mu, \Sigma) = \mathcal{N}_{\mathcal{S}}(\mu, \Sigma)$
- Quand $\Sigma^{-1} \to \mathbf{0}_{D \times D}$, $\mathcal{A}(\alpha, \mu, \Sigma) \to \mathcal{D}(\alpha, \dots, \alpha)$ en distribution.

Remarques

- Quand $\alpha = 0$, $\mathcal{A}(\alpha, \mu, \Sigma) = \mathcal{N}_{\mathcal{S}}(\mu, \Sigma)$
- Quand $\Sigma^{-1} \to \mathbf{0}_{D\times D}$, $\mathcal{A}(\alpha, \mu, \Sigma) \to \mathcal{D}(\alpha, \dots, \alpha)$ en distribution.

- Les paramètres μ et Σ ne **sont pas** la moyenne et la variance de $\mathcal{A}(\alpha, \mu, \Sigma)$
- Ils s'en rapprochent lorsque α tend vers 0.
- Il existe une autre paramétrisation, plus générale, de la densité (voir poly).

Remarques

- Quand $\alpha = 0$, $\mathcal{A}(\alpha, \mu, \Sigma) = \mathcal{N}_{\mathcal{S}}(\mu, \Sigma)$
- Quand $\Sigma^{-1} \to \mathbf{0}_{D\times D}$, $\mathcal{A}(\alpha, \mu, \Sigma) \to \mathcal{D}(\alpha, \dots, \alpha)$ en distribution.

- Les paramètres μ et Σ ne **sont pas** la moyenne et la variance de $\mathcal{A}(\alpha, \mu, \Sigma)$
- Ils s'en rapprochent lorsque α tend vers 0.
- Il existe une autre paramétrisation, plus générale, de la densité (voir poly).

La simulation est compliquée mais il existe des routines numériques dans `compositions::rAitchison()`

Illustrations ($\alpha = 1$ et $\alpha = 5$)

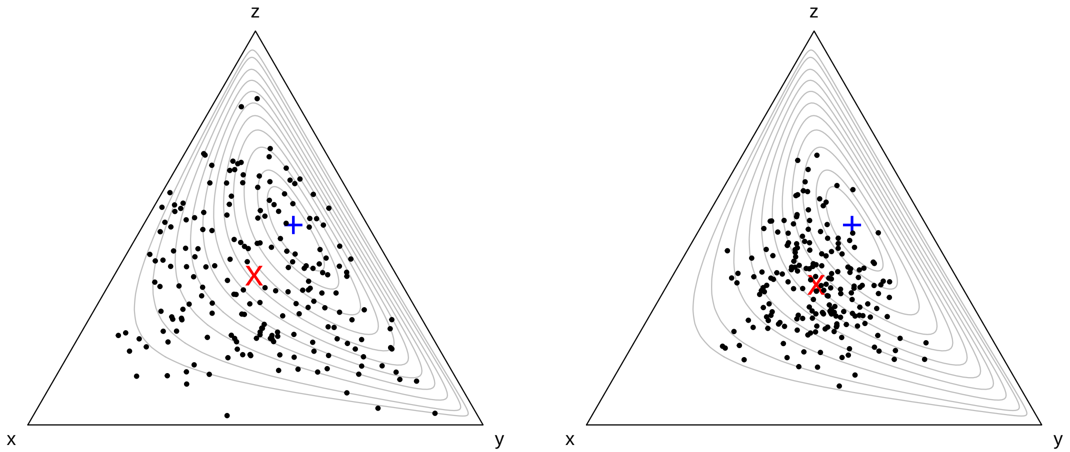

Figure 10: Loi de Aitchison de mêmes paramètres de (quasi-)moyenne et (quasi-)variance avec $\alpha = 1$ (gauche) et $\alpha = 5$ (droite). La moyenne empirique (x) est très différente de μ (+). Les isocontours sont ceux de la gaussienne.

Illustrations ($\alpha = 1$ et $\alpha = 0.3$)

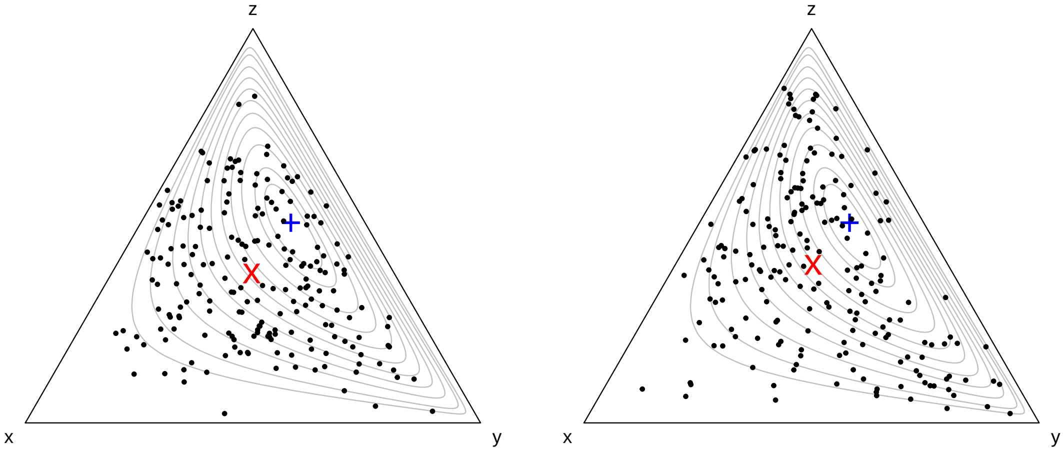

Figure 11: Loi de Aitchison de mêmes paramètres de (quasi-)moyenne et (quasi-)variance avec $\alpha = 1$ (gauche) et $\alpha = 0.3$ (droite). La moyenne empirique (x) est très différente de μ (+). Les isocontours sont ceux de la gaussienne.

Références

Aitchison, J. 1985. "A General Class of Distributions on the Simplex." *Journal of the Royal Statistical Society: Series B (Methodological)* 47 (1): 136–46. https://doi.org/https://doi.org/10.1111/j.2517-6161.1985.tb01341.x.

———. 1986. *The Statistical Analysis of Compositional Data (Monographs on Statistics and Applied Probability)*. Chapman & Hall Ltd., London, (Reprinted in 2003 with additional material by The Blackburn Press).

Figueras, Glòria Mateu. 2003. "Models de Distribució Sobre El śimplex." PhD thesis, Universitat Politècnica de Catalunya (UPC).

Van den Boogaart, K Gerald, and Raimon Tolosana-Delgado. 2013. *Analyzing Compositional Data with r*. Vol. 122. Springer.