CHARLET Antoine, DJITE Mouhamadou, MARLIER Laura

MASSABA DOUMBIA Awa, MONIN Louis

ING1-GM3

Projet Mathématiques & Informatique

Rapport n°3 : Projet final

Projet encadré par

BOURHATTAS Abderrahim

Professeur à CY-Tech

CY TECH – Cergy Paris Université

Année 2022-2023

Table des matières

Introduction

Dans le cadre de notre première année de cycle ingénieur en génie mathématiques à CY-Tech, nous avons reçu un projet visant à mettre en application nos différentes méthodes étudiées en cours, ainsi que de mettre en pratique une démarche rigoureuse d'analyse et de programmation d'un problème. Ce travail a eu pour objectif de développer nos compétences en analyse de données, en programmation ainsi que le travail d'équipe.

Notre équipe étant composé de CHARLET Antoine, DJITE Mouhamadou, MARLIER Laura, MASSABA DOUMBIA Awa, MONIN Louis. Nous avons sélectionné le thème de l'analyse exploratoire descriptive des données fonctionnelles.

Les statistiques descriptives et exploratoires sont des techniques courantes utilisées pour analyser les données fonctionnelles. Les données fonctionnelles sont des données qui sont mesurées sur une fonction continue, plutôt que sur des variables discrètes.

Les statistiques descriptives sont utilisées pour résumer les caractéristiques importantes des données fonctionnelles, telles que la tendance centrale, la dispersion et la forme de la distribution. La moyenne, la médiane et le mode sont des exemples de mesures de tendance centrale, tandis que l'écart-type et la variance sont des mesures de dispersion. Les graphiques tels que les histogrammes et les graphiques en boîte peuvent également être utilisés pour visualiser les données fonctionnelles.

Les statistiques exploratoires, quant à elles, sont utilisées pour découvrir des relations entre différentes variables fonctionnelles. Les techniques courantes utilisées pour l'analyse exploratoire des données fonctionnelles comprennent l'analyse de corrélation fonctionnelle, l'analyse en composantes principales fonctionnelles et la régression fonctionnelle.

En somme, les statistiques descriptives et exploratoires pour les données fonctionnelles sont des outils essentiels pour comprendre les caractéristiques et les relations entre les données fonctionnelles. Ces techniques peuvent aider les chercheurs à mieux comprendre les modèles de données fonctionnelles et à les utiliser pour prendre des décisions éclairées.

Ces éléments de recherche peuvent nous mener à nous demander si les problèmes statistiques qui peuvent se poser sont les mêmes qu'en statistiques multivariées classiques ? Pour parvenir à répondre à cette problématique, nous nous sommes penchés sur un jeu de données dénommé MontrealTemps (fda) qui regroupe la température notée T en degré Celsius à Montréal, au Canada, chaque jour de 1961 à 1994. Ainsi ce jeu de données possède le format suivant, un tableau numérique avec des dimnames=list(1961 :1994,names(dayOfYear)).

Pourquoi privilégier les données fonctionnelles ?

Cela dépend dans un premier temps de la nature des données. Les données fonctionnelles représentent généralement des mesures continues sur un domaine continu, ce qui peut être plus approprié pour modéliser des phénomènes qui évoluent dans le temps ou dans l'espace de manière continue. Les variables aléatoires classiques sont généralement discrètes, ou prennent des valeurs spécifiques. Cela ne peut être adapté à certains types de données tels que des signaux continus, des courbes de croissance ou des trajectoires temporelles.

De plus, les données fonctionnelles sont souvent utilisées pour modéliser des comportements complexes et dynamiques qui sont difficiles à capturer avec des variables aléatoires classiques. Par exemple, dans le cas de signaux audio, de données biomédicales ou de trajectoires de mouvement, les données fonctionnelles peuvent permettre de représenter de manière plus précise les variations continues et subtiles des données, tandis que des variables aléatoires classiques pourraient ne pas être en mesure de capter ces nuances.

En effet, dans notre cas d'étude des tendances temporelles, c'est-à-dire des évolutions dans le temps, les données fonctionnelles sont particulièrement adaptées pour les modéliser. Par exemple, les données fonctionnelles peuvent être utilisées pour modéliser la croissance d'une population, l'évolution de la température sur une période de temps ou la fluctuation des prix d'un actif financier au fil du temps. Les variables aléatoires classiques peuvent être moins appropriées pour modéliser ces types de tendances temporelles.

Enfin, les données fonctionnelles sont souvent plus faciles à interpréter et à visualiser. En effet, elles permettent de représenter graphiquement des trajectoires continues dans un espace fonctionnel. Cela peut faciliter la compréhension et l'interprétation des modèles statistiques basés sur des données fonctionnelles mais aussi la visualisation des résultats pour une communication plus claire.

Peut-on réaliser des statistiques avec les données fonctionnelles ?

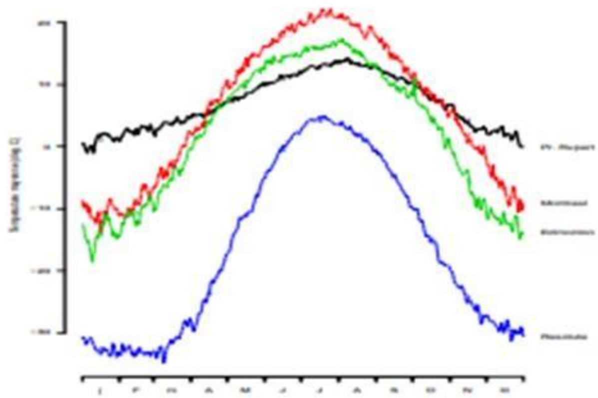
Les statistiques fonctionnelles permettent de développer des méthodes et des outils statistiques spécifiques pour traiter les particularités des données fonctionnelles, notamment leur nature continue, leur dimension infinie et leur comportement dynamique dans le temps ou dans l'espace. Les méthodes statistiques fonctionnelles sont souvent utilisées pour analyser des données provenant de domaines tels que la biologie, la médecine, la finance, la météorologie, l'économie, l'ingénierie, l'audio ou la vision par ordinateur, où les données fonctionnelles sont couramment rencontrées.

Les statistiques fonctionnelles peuvent inclure :

- Des méthodes de régression fonctionnelle pour modéliser les relations entre des variables fonctionnelles,
- Des méthodes d'analyse de la variance fonctionnelle pour comparer des groupes de fonctions,
- Des méthodes de classification fonctionnelle pour la classification de fonctions,
- Des méthodes de densité fonctionnelle pour l'estimation de la densité de probabilité des données fonctionnelles, et bien d'autres.

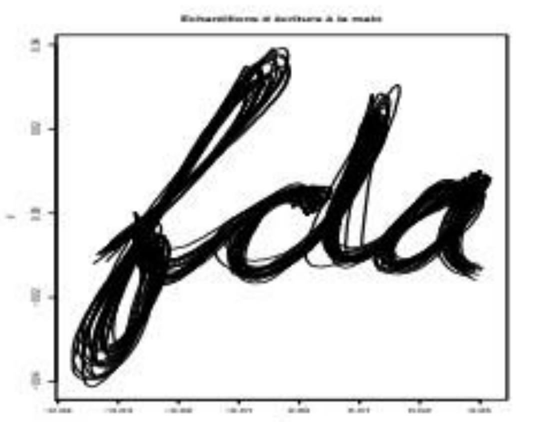
Les statistiques fonctionnelles peuvent également inclure des techniques de visualisation et d'interprétation spécifiques pour les données fonctionnelles, telles que les représentations graphiques de fonctions, les tracés de courbes de croissance et les cartes de chaleur fonctionnelles.

Les données longitudinales sont des courbes de températures relevées en un point donné à différents instants, suivi d'une population dans le temps relativement à un phénomène mesurable, courbe de croissance d'un individu.

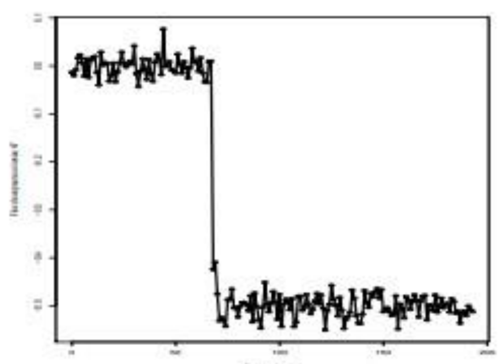


Courbes de température.

Les données multivariées ou spatiales permettent le suivi des mouvements de points à travers le plan ou de l'espace, flux de vapeur dans une même colonne en fonction du temps



20 tracés
(unité de l'axe des abscisses : cm.)

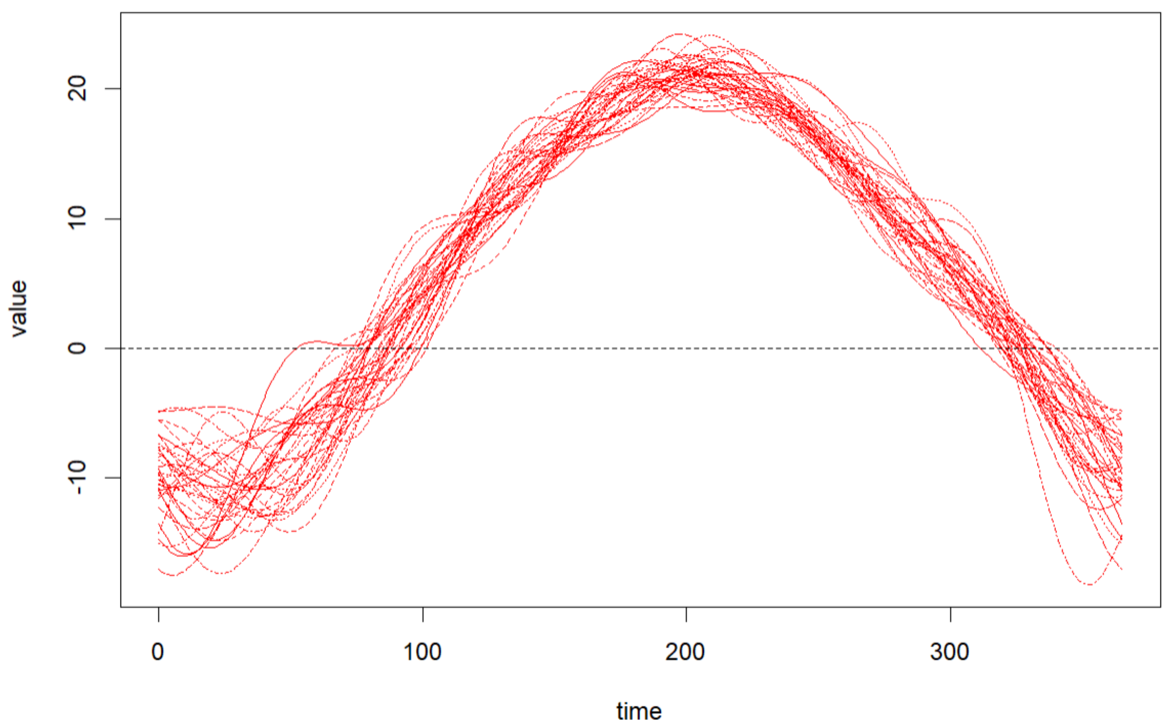


flux de vapeur
dans cette même colonne
en fonction du temps

Analyse théorique

Etude descriptive préliminaire

Notre jeu de données est composé de 34 courbes représentant la date du jour à une année précise en fonction de la température en °C. Le graphe ci-contre présente notre dataset de manière graphique.



A la suite de cette représentation, nous nous sommes orientés vers une étude préliminaire telles que la moyenne, la médiane, la variance ainsi que la covariance.

Moyenne

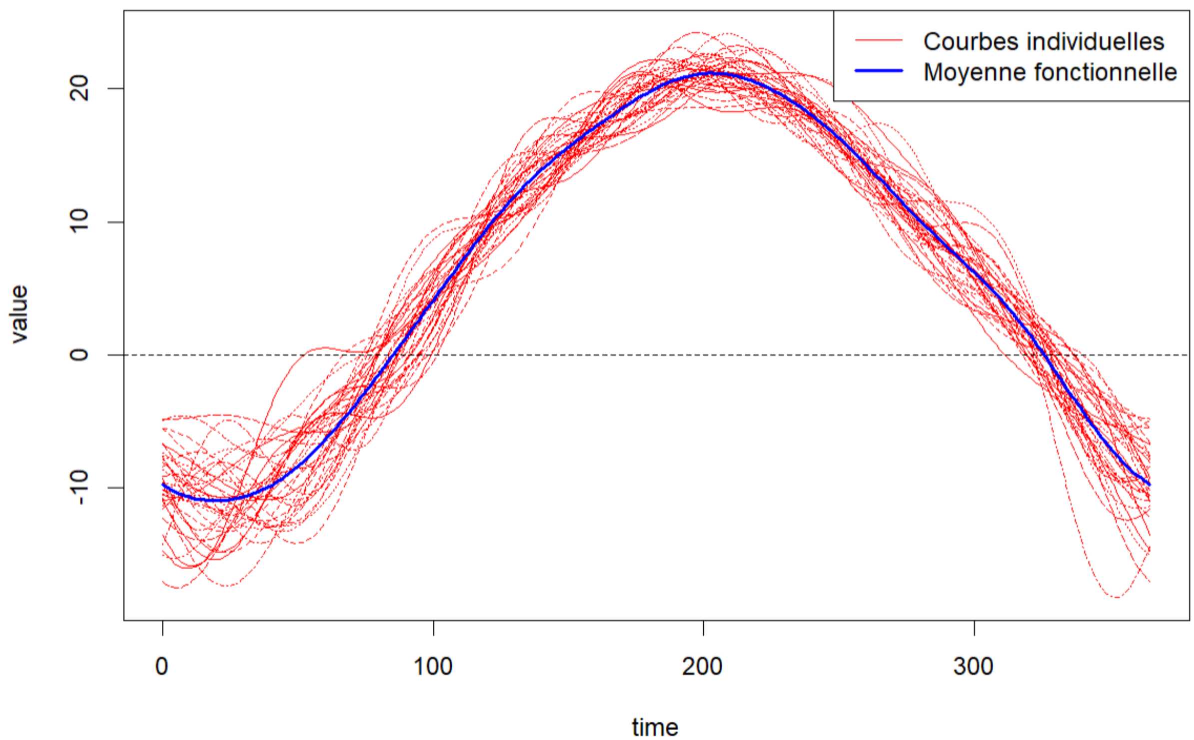
Définition : Fonction moyenne empirique :

$$X_n : T \rightarrow \mathbb{R}$$

$$t \rightarrow X_n(t) = \frac{1}{n}\sum_{i}^{n} X_i(t)$$

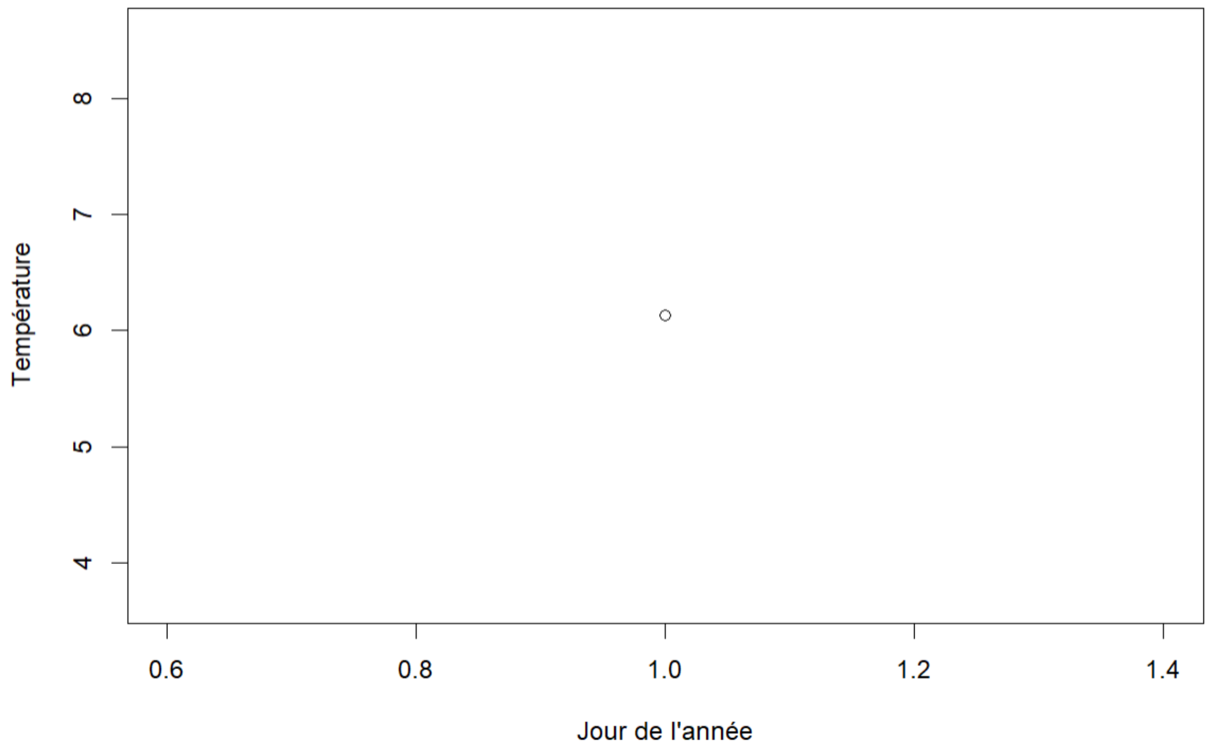
La notion de moyenne pour des données fonctionnelles est utilisée dans le domaine de l'analyse de données fonctionnelles, qui traite de la manipulation et de l'analyse de variables qui varient de manière continue sur un intervalle donné, plutôt que de manière discrète comme les variables traditionnelles.

Lorsqu'il s'agit de données fonctionnelles, la moyenne représente essentiellement une fonction qui résume le comportement moyen de l'ensemble des fonctions observées. La moyenne fonctionnelle est souvent utilisée pour obtenir une représentation globale et caractéristique d'un ensemble de courbes ou de fonctions.



Mode

Le dataset MontrealTemp représente les données de température à Montréal sur une période donnée. Pour représenter le mode fonctionnel dans ce contexte, nous pouvons estimer la trajectoire de température qui est la plus fréquente ou dominante dans l'ensemble des données (voir code en annexe). De ce fait nous obtenons une valeur d'environ 6.06°C. Ce qui semble cohérent avec les températures glaciales que peut connaitre le climat au Canada.



Variance

- ***fonction variance empirique :***

$$
\begin{array}{rcl}
Var_n : T & \longrightarrow & \mathbb{R}, \\
t & \longmapsto & Var_n(t) = \frac{1}{n} \sum_{i=1}^{n} \left(X_i(t) - \bar{X}_n(t) \right)^2 .
\end{array}
$$

On définit également l'écart-type comme étant la racine carrée de cette fonction variance.

La variance pour des données fonctionnelles est une mesure de dispersion qui quantifie la variabilité des fonctions observées dans un ensemble de données fonctionnelles. Contrairement à la variance pour des données numériques ou discrètes, la variance pour des données fonctionnelles prend en compte les variations continues le long d'une dimension fonctionnelle.

Pour calculer la variance fonctionnelle, on peut utiliser la méthode de l'analyse des composantes principales fonctionnelles (Functional Principal Component Analysis - FPCA). Cette méthode permet de décomposer les fonctions observées en une série de fonctions de base orthogonales (fonctions propres) et d'identifier les coefficients correspondants (scores). La variance fonctionnelle est ensuite calculée en prenant en compte la contribution de chaque fonction propre à la variabilité totale.

Une autre approche courante est d'utiliser la notion de covariance fonctionnelle. La covariance fonctionnelle mesure la similarité entre deux fonctions à différents points le long de la dimension fonctionnelle. En utilisant la covariance fonctionnelle, on peut calculer la variance fonctionnelle en considérant la variation entre les fonctions et la variation interne de chaque fonction.

Il est important de noter que la définition précise de la variance fonctionnelle peut dépendre du contexte spécifique de l'analyse de données fonctionnelles et des méthodes utilisées. Ces approches permettent de quantifier la dispersion et la variabilité des fonctions dans un ensemble de données fonctionnelles, ce qui est utile pour l'interprétation et l'analyse statistique.

On peut définir au moins deux notions d'espérance pour un processus stochastique. Une première manière rigoureuse et générale consiste à introduire l'intégrale de Bochner qui gé néralise l'intégrale de Lebesgue pour des fonctions à valeurs dans un espace de Banach (voir Dunford et Schwartz 1958).

Si $\mathbb{E}[\|X\|] < \infty$*, on peut définir l'**espérance** de* X *comme*

$$\mathbb{E}[X] = \int_\Omega X(\omega) d\mathbb{P}(\omega),$$

*où l'intégrale est l'**intégrale de Bochner** de* X*. On définit ainsi un élément de* $\mathcal{F}$*. On dira que* X *est centrée lorsque* $\mathbb{E}[X] = 0$*.*

Cette espérance a des propriétés "classiques" : par exemple, $\|\mathbb{E}[X]\| \leq \mathbb{E}[\|X\|]$, ou $\|\mathbb{E}[X]\|^2 \leq \mathbb{E}[\|X\|^2]$. Lorsque $\mathcal{F}$ est un espace de Hilbert, de produit scalaire $\langle \cdot, \cdot \rangle$, on a en fait plus généralement $\mathbb{E}[\langle X, f \rangle] = \langle \mathbb{E}[X], f \rangle$.

Une autre possibilité pour définir l'espérance d'une variable fonctionnelle consiste à prendre point par point l'espérance de X : on se contentera de la définition (légèrement imprécise) suivante.

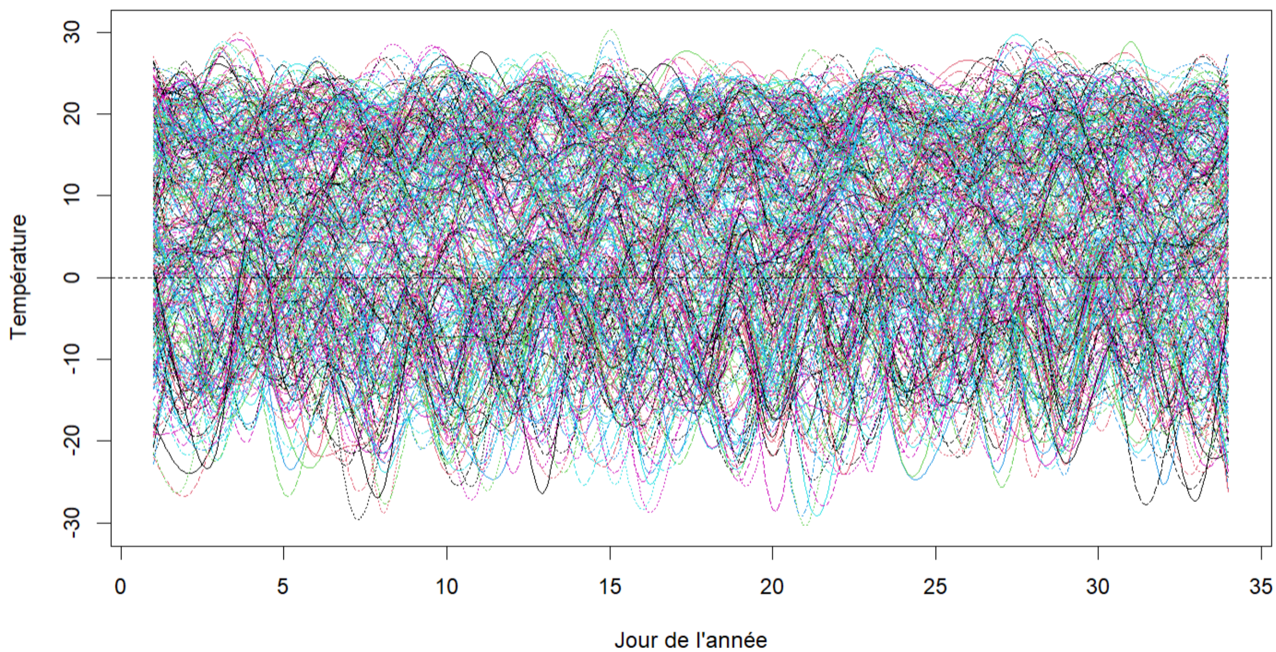
*Lorsque cela a un sens, on peut définir l'**espérance** de* X *comme la fonction*

non aléatoire $(\mathbb{E}[X])(\cdot)$ *(élément de* $\mathcal{F}$*) définie point par point par*

$$(\mathbb{E}[X])(t) = \mathbb{E}[X(t)] = \int_\Omega X(t, \omega) d\mathbb{P}(\omega),$$

Ces deux notions d'espérance coïncident souvent. Par exemple, si $\mathcal{F} = L^2(T)$, l'espérance point à point de X existe et définit bien un élément de $L^2(T)$ dès que $\mathbb{E}[\|X\|] < \infty$, et on a égalité presque partout des deux notions d'espérance dans ce cas.

Dans notre cas présent (voir code annexe), nous avons obtenu le graphique ci-contre, où nous avons représenter la variance de chaque mois de chaque année.



Covariance

Dans le cadre multivarié, la covariance a une forme matricielle : l'extension naturelle à la dimension infinie est donc un opérateur. On se restreint à présent au cas où F est un espace de Hilbert.
Supposons que $\mathbb{E}[\|X\|^2] < \infty$. *L'**opérateur de covariance** de* X *est défini*

par

$$\Gamma \,:\, f \in \mathcal{F} \longmapsto \Gamma f = \mathbb{E}[\langle X - \mathbb{E}[X], f\rangle (X - \mathbb{E}[X])].$$

Il est à valeurs dans $\mathcal{F}$.

Dans le cas où X est centrée, on a la réécriture plus simple, $\Gamma \,:\, f \in \mathcal{F} \longmapsto \mathbb{E}[\langle X, f\rangle X]$. Notons qu'une telle définition est en fait possible également dans le cas où $\mathcal{F}$ est seulement un espace de Banach. On remplace alors le produit scalaire par le *crochet de dualité*, Γ appliquant alors $\mathcal{F}^*$ dual de $\mathcal{F}$ dans $\mathcal{F}$. Mais les propriétés qui suivent vont justifier qu'on se borne au cas hilbertien. Dans tous les cas, Γ est un *opérateur linéaire* (ceci découle de la linéarité du produit scalaire), et *continu*, puisque $\|\Gamma f\| \leq \mathbb{E}[\|X\|^2]\|f\|$ (en utilisant l'inégalité de Cauchy-Schwarz). On peut donc voir l'opérateur de covariance comme une matrice de dimension infinie, extension directe de la notion de matrice de variance-covariance d'un vecteur aléatoire.

On a, dans le cas hilbertien, le résultat suivant, que l'on pourra omettre si l'on n'est pas familier de la théorie des opérateurs, pour n'en retenir que les conséquences (voir section suivante).

Supposons que X est une variable aléatoire à valeurs dans un espace de

Hilbert séparable $\mathcal{F}$ telle que $\mathbb{E}[\|X\|^2] < \infty$. Alors, son opérateur de covariance Γ est

(i) ***autoadjoint*** *: $\forall f, g \in \mathcal{F}$, $\langle \Gamma f, g\rangle = \langle f, \Gamma g\rangle$,*

(ii) ***de Hilbert-Schmidt*** *: il existe une base hilbertienne $(\varphi_j)_{j\in\mathbb{N}}$ telle que $\sum_{j\in\mathbb{N}} |\Gamma\varphi_j|^2 < \infty$.*

On en déduit,

Avec les hypothèses et notations de la proposition précédente, l'opérateur de

covariance Γ de X est ainsi

(i) ***compact*** *(l'image de la boule unité de $\mathcal{F}$ par Γ est relativement compacte),*

(ii) ***diagonalisable en base orthonormée****.*

Le premier point est une conséquence du caractère Hilbert-Schmidt de Γ. Le second point provient du résultat fondamental de décomposition des opérateurs autoadjoints compacts. Nous tirerons les conséquences de cette propriété à la Section 2.2.3 ci-dessous. Enfin, on notera que si Γ est de rang fini, alors X appartient en fait un espace de dimension finie.

Une autre quantité que l'opérateur de covariance est fréquemment utilisée en théorie des processus stochastiques. Pour simplifier (encore), on suppose que $\mathcal{F}$ est l'espace $L^2([0,1])$.

Soit X une variable aléatoire à valeurs dans $L^2([0,1])$. La ***fonction de***

covariance *de X est l'application*

$$C : (s,t) \in [0,1]^2 \longmapsto C(s,t) = \mathbb{E}[X(t)X(s)].$$

Le lien entre fonction de covariance et l'opérateur de covariance est donné par le résultat suivant.

Soit X une variable aléatoire à valeurs dans $L^2([0,1])$, d'opérateur de

covariance Γ et de fonction de covariance C. Alors,

$$\forall f \in L^2([0,1]),\ \forall t \in [0,1],\ (\Gamma f)(t) = \int_{[0,1]} C(s,t) f(s) ds.$$

Ainsi, Γ est un ***opérateur à noyau****, de noyau la fonction de covariance.*

Matrice de covariance

La matrice de covariance d'un dataset de données fonctionnelles est une matrice carrée qui mesure les relations de covariance entre les différentes variables fonctionnelles du dataset.

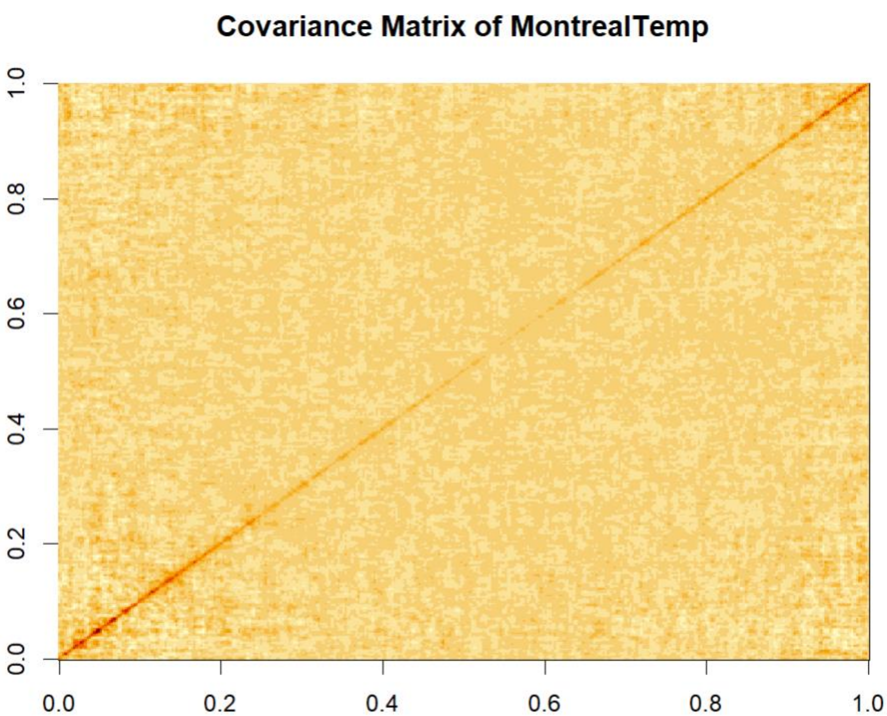
Pour un dataset de données fonctionnelles, chaque observation est une fonction qui varie en fonction d'une variable indépendante (par exemple, le temps). Chaque fonction est représentée par un ensemble de valeurs ou de points dans un espace fonctionnel.

La matrice de covariance est construite en calculant les covariances entre paires de fonctions dans le dataset. La covariance mesure la variation conjointe entre deux fonctions à travers leurs valeurs respectives à différents points de la variable indépendante. Une valeur de covariance positive indique une variation similaire entre les fonctions, tandis qu'une valeur négative indique une variation opposée.

La matrice de covariance est carrée, de dimensions p x p, où p est le nombre de variables fonctionnelles dans le dataset. Les éléments de la matrice de covariance sont les covariances entre les différentes paires de variables fonctionnelles. La diagonale de la matrice contient les variances des variables fonctionnelles individuelles.

La matrice de covariance est une mesure utile pour analyser les relations de covariance entre les variables fonctionnelles, identifier les tendances communes et les dépendances entre les fonctions, et effectuer des analyses multivariées sur les données fonctionnelles.

Nous avons donc étudié et générer la matrice de covariance de notre dataset pour voir si elle respecte correctement l'ensemble des points de définition de la matrice de covariance (voir code en annexe).



Ainsi, après l'obtention de notre matrice de covariance nous avons pu constater que notre matrice de covariance liée à notre dataset respectait correctement les éléments de la définition de la matrice de covariance.

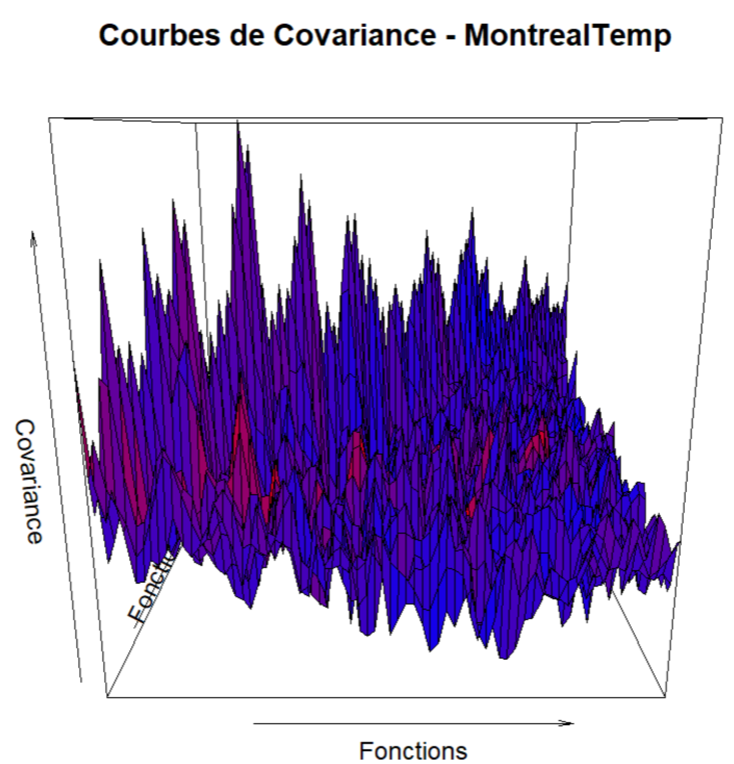
Covariance en 3D

La covariance en 3D d'un dataset de données fonctionnelles représente la mesure de la covariation entre les différentes variables fonctionnelles dans un espace tridimensionnel. Elle permet de quantifier les relations linéaires entre les fonctions dans les trois dimensions.

Pour calculer la covariance en 3D, chaque paire de fonctions est évaluée pour mesurer leur degré de similitude ou de corrélation dans les trois dimensions de l'espace fonctionnel.

La covariance en 3D est une matrice tridimensionnelle symétrique. Les éléments diagonaux de cette matrice représentent la variance des fonctions individuelles le long des trois dimensions, tandis que les éléments non diagonaux représentent la covariance entre les paires de fonctions dans les trois dimensions.

La covariance en 3D est utile pour analyser les relations et les structures complexes qui peuvent exister entre les fonctions dans un espace tridimensionnel. Elle peut être utilisée dans diverses techniques d'analyse statistique, telles que l'Analyse en Composantes Principales (ACP) fonctionnelle en 3D, pour extraire des informations significatives et identifier les principales sources de variation dans les données fonctionnelles.



Les pics dans la courbe de covariance indiquent des relations fortes ou des covariations positives entre la température et la date, tandis que les creux suggèrent des covariations négatives ou des relations inverses. On remarque alors que les données fonctionnelles de notre dataset on des covariations positives ce qui impliquent des relations fortes entre la température et la date.

Corrélations

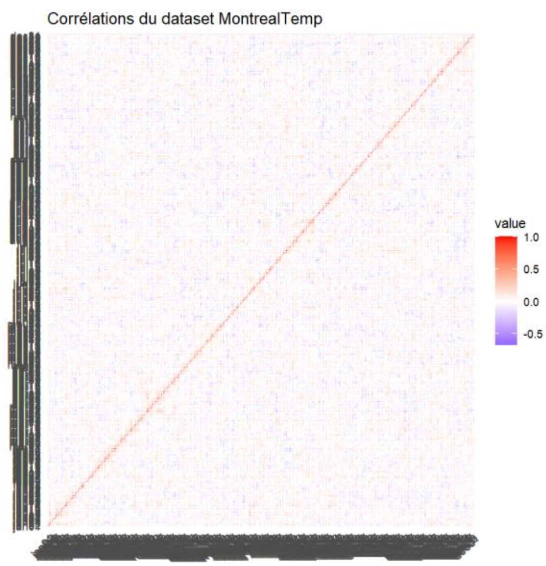
Les corrélations entre les données fonctionnelles du dataset "MontrealTemp" mesurent la relation linéaire entre les différentes variables fonctionnelles représentant les températures à Montréal. Ces corrélations indiquent dans quelle mesure les variations d'une variable fonctionnelle sont associées aux variations des autres variables fonctionnelles.

Les corrélations entre les données fonctionnelles peuvent être calculées à l'aide de mesures de corrélation appropriées pour les données fonctionnelles. Une mesure couramment utilisée est la corrélation fonctionnelle, qui évalue la similitude des motifs temporels entre les variables fonctionnelles.

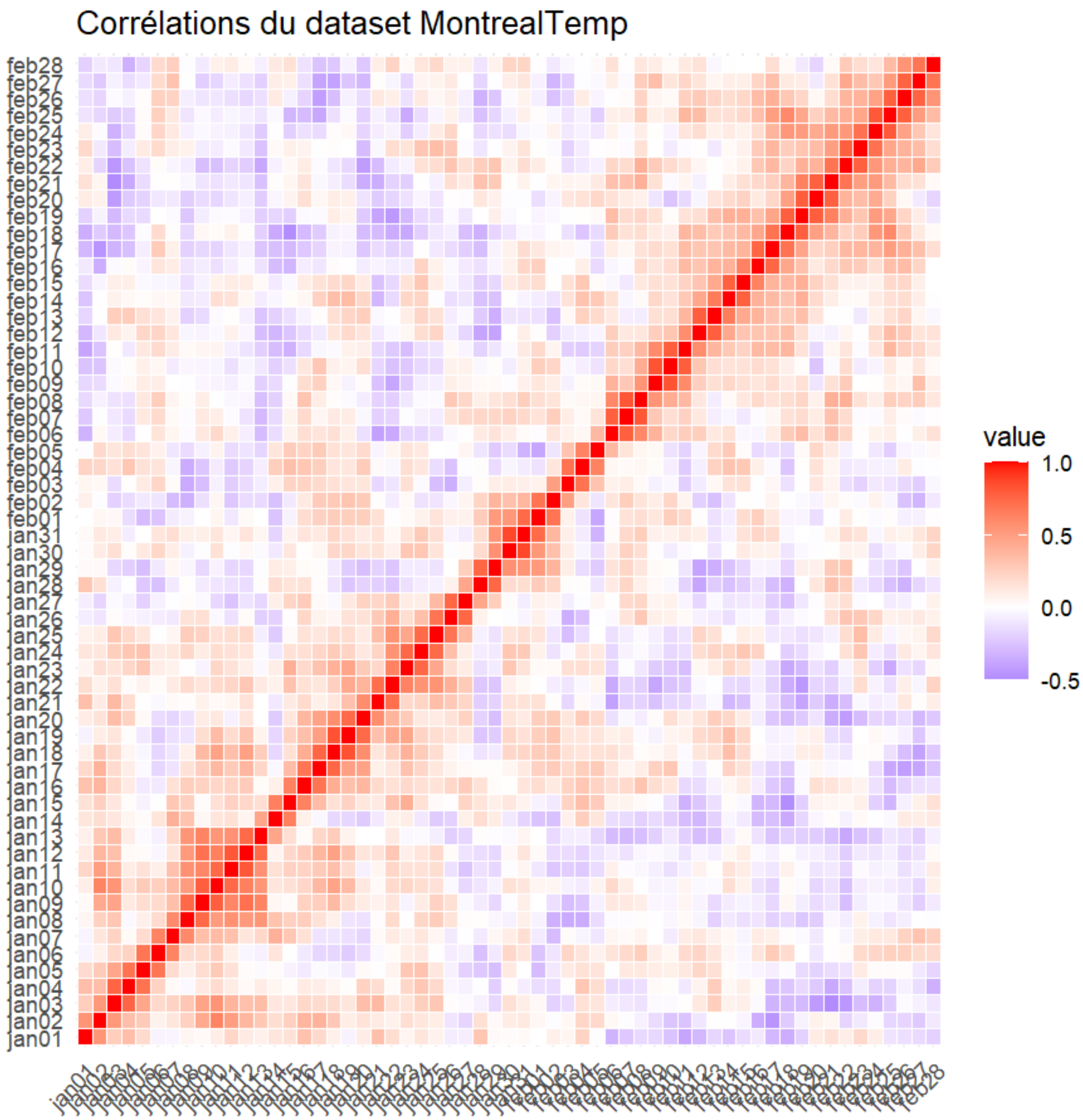
En calculant les corrélations entre les données fonctionnelles du dataset MontrealTemp, vous pouvez déterminer les relations de similarité ou de dissimilarité entre les différentes variables fonctionnelles représentant les températures à différents endroits de Montréal.

Ces corrélations peuvent être utiles pour identifier les zones géographiques ayant des profils de température similaires, détecter les tendances ou les motifs de variation communs entre les variables fonctionnelles, ou encore pour explorer les liens entre les conditions météorologiques dans différentes régions de la ville.

Il convient de noter que l'interprétation des corrélations entre les données fonctionnelles dépendra du contexte spécifique de l'analyse et des objectifs de l'étude. Dans notre cas, en prenant en compte la totalité du dataset, nous avons obtenue le graphique suivant (voir code en annexe).



Nous avons en abscisse les dates ainsi qu’en ordonnée. Or, il nous est impossible de lire graphiquement à cause du trop grand nombre de variables. Pour ce fait, nous avons réduit le nombre de variables. Nous avons donc obtenu le graphique ci-contre (voir code en annexe).



On peut donc facilement distinguer les corrélations entre les différentes variables. On remarque aisément que les corrélations sont positives entre les mois de janvier des différentes années mais que ces corrélations deviennent négatives entre le mois de janvier et le mois de février ce qui s’étend au même résultat sur tous les mois de l’année sur différentes années.

Analyse des méthodes employés

Logiciel

Pour ce projet, nous avons utilisé le langage R à l'aide du logiciel RStudio. R étant un outil de choix pour les études exploratoires et descriptives de données fonctionnelles grâce à sa richesse de packages, sa flexibilité et sa capacité à effectuer des analyses statistiques sophistiquées sur ce type de données.

ACPF

L'Analyse des Composantes Principales Fonctionnelles (ACPF) est une méthode statistique utilisée pour l'analyse de données fonctionnelles. Elle permet de réduire la dimensionnalité des données fonctionnelles et d'extraire les principales sources de variation présentes dans l'ensemble des fonctions observées.

L'ACPF repose sur le principe de décomposer les fonctions observées en une série de fonctions de base orthogonales appelées "fonctions propres" ou "fonctions principales". Ces fonctions propres représentent les modes de variation dominants dans les données fonctionnelles.

Le processus de l'ACPF consiste généralement en plusieurs étapes :

1. Standardisation : Les données fonctionnelles sont souvent standardisées pour rendre les variables comparables et les mettre à la même échelle.

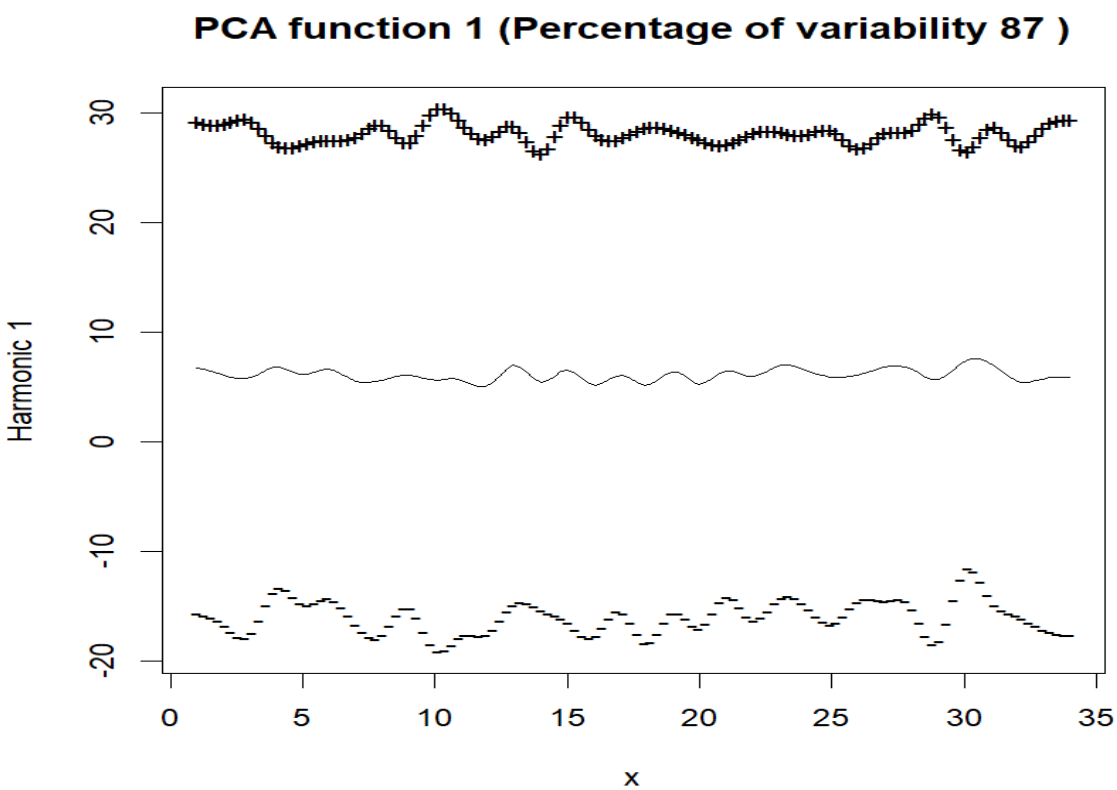
2. Décomposition : Les fonctions sont décomposées en une série de fonctions propres en utilisant des méthodes telles que l'analyse en composantes principales (PCA) fonctionnelle ou l'analyse spectrale.

3. Sélection des composantes principales : Les fonctions propres sont ordonnées en fonction de leur contribution à la variabilité totale. Les premières composantes principales représentent les modes de variation les plus importants.

4. Interprétation : Les composantes principales sont interprétées en analysant les fonctions propres correspondantes. Elles peuvent représenter des motifs, des tendances ou des caractéristiques spécifiques des données fonctionnelles.

5. Réduction de dimension : Les composantes principales les plus significatives peuvent être sélectionnées pour réduire la dimensionnalité des données fonctionnelles, permettant ainsi une représentation plus concise et une simplification de l'analyse.

Graphiquement on a représenté la première composante principale fonctionnelle comme fonction du temps (voir code en annexe)



C’est une représentation graphique de la première CPF en fonction du temps. Chaque CPF est, avant tout, une fonction (continue) du temps, considéré comme un vecteur dans un certain espace vectoriel des fonction F (intégrable en L^2) définies sur un certain intervalle de temps.

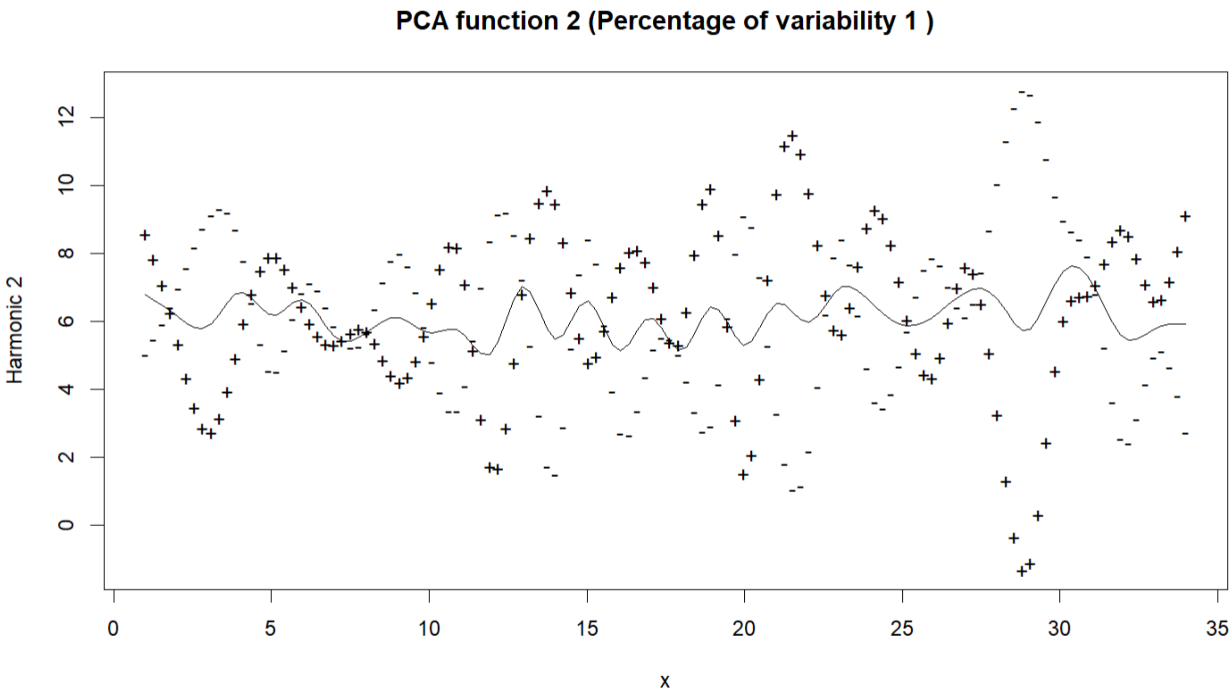
Valeurs propres

Les valeurs propres dans une Analyse en Composantes Principales (ACP) fonctionnelle sont des mesures de la variance expliquée par chaque composante principale. Dans une ACP fonctionnelle, les données sont représentées par des fonctions continues plutôt que par des vecteurs numériques.

Les valeurs propres sont calculées en effectuant une décomposition spectrale de la matrice de covariance fonctionnelle des données. Chaque valeur propre indique la proportion de variance totale expliquée par la composante principale correspondante. Plus une valeur propre est élevée, plus la composante principale correspondante est importante pour la variation des données.

En résumé, les valeurs propres dans une ACP fonctionnelle mesurent l'importance relative de chaque composante principale dans la représentation des variations dans les données fonctionnelles.

Dans notre dataset, nous avons extrait les valeurs propres de notre jeu de données (voir code en annexe) et nous avons obtenu le graphique ci-contre.



Conclusion

En conclusion, l'analyse des données fonctionnelles de la dataset MontrealTemp a fourni des informations précieuses sur les tendances climatiques à Montréal. La dataset comprend des mesures de température sur une période prolongée, ce qui permet d'obtenir une vision globale des variations saisonnières et des changements à long terme.

Au cours de notre analyse, nous avons identifié plusieurs points clés :

-Tendances saisonnières : Les données ont révélé des variations saisonnières régulières de la température à Montréal. Nous avons observé des étés relativement chauds avec des températures moyennes élevées, tandis que les hivers ont tendance à être froids avec des températures moyennes plus basses. Ces tendances saisonnières sont conformes aux attentes générales pour cette région.

-Changements à long terme : Une analyse plus approfondie a révélé des indications de changements à long terme dans les températures à Montréal. Au fil des décennies, nous avons constaté une légère augmentation de la température moyenne annuelle. Bien que cette augmentation puisse sembler modeste, elle est cohérente avec les tendances mondiales du réchauffement climatique. Ces résultats soulignent la nécessité d'une surveillance continue des changements climatiques dans la région.

-Extrêmes climatiques : Les données fonctionnelles ont également permis d'identifier les événements climatiques extrêmes à Montréal. Nous avons observé des températures maximales exceptionnellement élevées pendant certains étés, ainsi que des températures minimales extrêmement basses pendant certains hivers. Ces extrêmes peuvent avoir des implications importantes pour les infrastructures, l'agriculture et la santé publique, soulignant l'importance de l'adaptation aux changements climatiques.

En résumé, l'analyse des données fonctionnelles de la dataset MontrealTemp a fourni des informations clés sur les tendances saisonnières, les changements à long terme et les extrêmes climatiques à Montréal. Ces résultats soulignent l'importance de la surveillance continue du climat et de la prise de mesures appropriées pour atténuer les effets du réchauffement climatique.