

Modélisation des Noyaux Spatio-Temporels Financiers

Mohamed Yahya

Superviseur : Prof. Shengrui Wang

Université de Sherbrooke

December 12, 2024

1 Introduction

1.1 Contexte

Dans le domaine de la finance quantitative, la prédiction précise des rendements d'actifs financiers demeure un défi majeur. L'utilisation de méthodes d'apprentissage automatique, en particulier la Régression par Processus Gaussien (GPR), offre une approche prometteuse pour modéliser ces séries temporelles complexes tout en quantifiant l'incertitude associée aux prédictions.

1.2 Problématique

La difficulté principale réside dans la capture efficace des dépendances temporelles et des relations entre différents actifs financiers. Les méthodes traditionnelles présentent souvent des limitations dans leur capacité à :

- Modéliser les relations non-linéaires complexes dans les données financières
- Quantifier l'incertitude des prédictions de manière fiable
- Capturer les dépendances temporelles à différentes échelles
- Prendre en compte les corrélations entre différents actifs

1.3 Objectifs

Ce projet vise à développer un cadre robuste de prédiction des rendements financiers en utilisant des modèles GPR avec différents noyaux. Les objectifs spécifiques sont :

- Développer et comparer différents noyaux pour la modélisation des séries temporelles financières
- Évaluer la performance de chaque noyau dans la capture des dépendances temporelles
- Étendre l'analyse à un portefeuille multi-actifs pour étudier les relations inter-actifs
- Fournir des estimations d'incertitude fiables pour la gestion des risques

1.4 Approche Méthodologique

Notre approche se décompose en deux parties principales :

1. **Analyse d'un actif unique** : Nous commençons par étudier la performance des différents noyaux sur un seul actif (IBM) pour établir une base de comparaison.
2. **Extension multi-actifs** : Nous étendons ensuite l'analyse à un portefeuille diversifié pour évaluer la capacité des modèles à capturer les relations inter-actifs.

2 Analyse pour un actif unique

2.1 Objectif

L'objectif de cette première partie est de prédire les rendements d'un actif unique, ici représenté par l'exemple d'IBM. Cette étape sert à tester l'efficacité des noyaux dans la capture des dépendances temporelles.

2.2 Étapes de la modélisation

2.2.1 Téléchargement et préparation des données

- Les données historiques de l'actif sont téléchargées depuis Yahoo Finance pour la période de septembre 2019 à septembre 2022. Le prix de clôture (*Close*) est utilisé pour calculer les rendements.
- Une méthode de fenêtre glissante (*rolling window*) est utilisée avec une taille de fenêtre $w = 5$:

$$\mathbf{x}_t = (P_{t-5}, P_{t-4}, \dots, P_{t-1}), \quad (1)$$

où P_t désigne le prix de clôture à la période t . La valeur cible correspond au prix de clôture suivant : $y_t = P_t$.

- Les données sont divisées en deux ensembles :
 - Ensemble d'entraînement (80% des données) : utilisé pour ajuster les paramètres des modèles.
 - Ensemble de test (20%) : utilisé pour évaluer les performances.

2.2.2 Normalisation des données

- Les données d'entrée (\mathbf{x}_t) et les cibles (y_t) sont normalisées à l'aide de `StandardScaler`.
- La normalisation est nécessaire pour améliorer la convergence des modèles et garantir que toutes les caractéristiques sont sur des échelles similaires.

2.2.3 Entraînement des modèles

- Les modèles *Gaussian Process Regression* (GPR) sont ajustés avec différents noyaux :
 - RBF
 - Matérn
 - Rational Quadratic
 - Combined (combinaison pondérée des trois précédents).
- Chaque noyau calcule la covariance entre deux séquences de rendements :

$$K(\mathbf{x}, \mathbf{x}') = \kappa_\theta(\mathbf{x}, \mathbf{x}'), \quad (2)$$

où θ désigne les hyperparamètres du noyau.

2.2.4 Prédictions et incertitudes

- Les prédictions sur l'ensemble de test sont accompagnées d'estimations d'incertitude fournies par le modèle GPR :

$$p(y_t | \mathbf{x}_t,) \sim \mathcal{N}(\mu_t, \sigma_t^2), \quad (3)$$

où μ_t est la prédiction et σ_t l'incertitude.

- Les prédictions et incertitudes sont dénormalisées pour permettre la comparaison avec les données réelles.

2.2.5 Évaluation des performances

- Les performances des modèles sont évaluées à l'aide de métriques :
 - RMSE : $\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$,
 - MAE : $\frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$,
 - MSE : $\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$,
 - Log-Marginal Likelihood (LML) : mesure l'adéquation des données au modèle.
- Les résultats sont visualisés sous forme de graphiques montrant :
 - Les prédictions (μ_t) et les bandes d'incertitude ($[\mu_t - 2\sigma_t, \mu_t + 2\sigma_t]$).
 - Les séries d'entraînement et de test pour comparaison.

2.3 Résultats et Analyse

2.3.1 Comparaison des métriques

| Kernel | RMSE | MAE | MSE | LML |
|-------------|-------------|-------------|-------------|--------------|
| RBF | 1.99 | 1.45 | 3.96 | 13.97 |
| Matérn | 2.03 | 1.52 | 4.14 | 22.32 |
| Rational Q. | 2.05 | 1.53 | 4.18 | 20.47 |
| Combined | 2.04 | 1.54 | 4.17 | 31.48 |

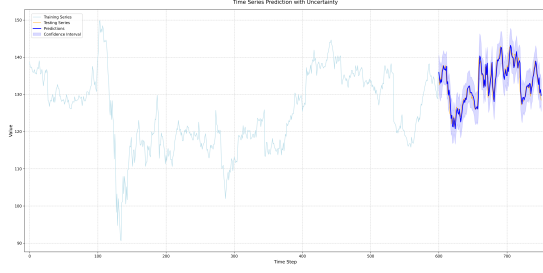
Table 1: Métriques de performance pour chaque modèle de noyau. Les meilleurs résultats sont en gras. Des valeurs plus faibles indiquent une meilleure performance pour toutes les métriques sauf le LML, où une valeur plus élevée est meilleure.

2.3.2 Résultats - Comparaison des différents noyaux

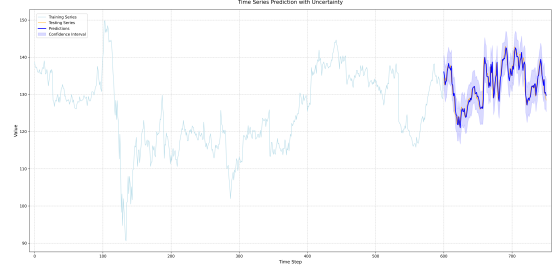
2.3.3 Analyse des résultats

Les résultats montrent que :

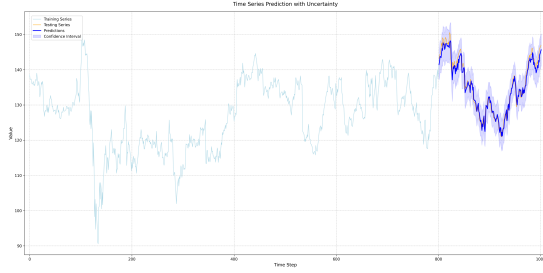
- Le noyau ****RBF**** excelle en précision pure avec les meilleures métriques d'erreur (RMSE : 1.99, MAE : 1.45).
- Le noyau ****Combined**** obtient le meilleur LML (31.48), indiquant une meilleure capture de la structure des données.



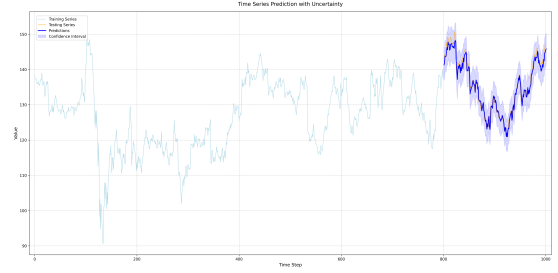
(a) Combined predictions



(b) RBF predictions



(c) Matérn predictions



(d) RQ predictions

Figure 1: Comparaison des performances des différents noyaux pour la prédiction.

2.4 Conclusion de la première partie

Cette étape a permis de valider l'efficacité des différents noyaux pour un actif unique. Les résultats obtenus servent de base pour l'extension à un portefeuille multi-actifs, où les relations inter-actifs seront analysées.

3 Analyse pour un portefeuille multi-actifs

3.1 Objectif

L'objectif de cette deuxième partie est d'étendre les concepts appliqués à un actif unique à un portefeuille multi-actifs. L'accent est mis sur l'analyse des relations inter-actifs et sur la prédiction des rendements d'un ensemble d'actifs.

3.2 Temps et données

Les points suivants décrivent systématiquement comment on prépare nos données.

- Soit un ensemble de points dans le temps tel que

$$= \{t_0, t_1, \dots, t_{T_r-1}, t_{T_r}, t_{T_r+1}, \dots, t_T\} \quad (4)$$

est défini étant une fréquence donnée¹.

¹Par exemple une fréquence mensuelle veut dire que chaque point $t \in$ désigne un mois donné. Dans ce cas, c'est le rendements du dernier jour ouvrable qui reflète le rendements d'un actif pour un mois donné.

- Additionnellement, suppose qu'on dispose d'un ensemble d'actifs (univers d'actifs)

$$= \{a_1, a_2, \dots, a_n\}, \quad (5)$$

avec chaque a_i dénote le symbole d'un actif donnée (e.g $a_j = \text{'META'}$).

- Soit $a \in$ un actif donné, on dénote par $P_{a,t}$ le prix (closing price) de l'actif a en point temporel $t \in$.
- Suppose que nous découpons notre ensemble de points temporels en deux parties, une partie avant T_r appelée ensemble de points temporels d'entraînement dénoté par r tel que

$$r = \{t_0, t_1, \dots, t_{T_r-1}, t_{T_r}\}. \quad (6)$$

- On appelle la deuxième partie l'ensemble de test dénoté par t_s tel que

$$s = \{t_{T_r+1}, t_{T_r+2}, \dots, t_{T-1}, t_T\}. \quad (7)$$

- On désigne par *une période* (lookback window) de taille w une suite de points temporels

$$\mathbf{t}_i = (t_{i-w}, t_{i-w+1}, \dots, t_{i-1}) \quad (8)$$

tirés de notre ensemble de points donné dans (4), et $w \leq i \leq T$ ².

- Soit \mathbf{t}_i une période de taille w , on désigne par $\mathbf{x}_{a,\mathbf{t}_i}$ la séquence de rendements observés pour un actif $a \in$ durant la période \mathbf{t}_i , ça veut dire

$$\mathbf{x}_{a,\mathbf{t}_i} = (R_{a,t_{i-w}}, R_{a,t_{i-w+1}}, \dots, R_{a,t_{i-1}}), \quad (9)$$

où R_{a,t_j} dénote le retour sur investissant dans l'actif a entre $t - 1$ et t :

$$R_{a,t_j} = \frac{P_{t_j} - P_{t_{j-1}}}{P_{t_{j-1}}} \quad (10)$$

On appelle une séquence de la sorte de (9) une séquence de rendements de l'actif a , et ce sont ces types de séquences qui vont servir comme input à nos kernels³.

- Soit t_i une période temporelle de taille w , on désigne par $y_{a,i}$ le rendements $y_{a,\mathbf{t}_i} = R_{a,t_i}$ pour décrire qu'après la séquence de rendements de l'actif $a \in$ observée durant la période \mathbf{t}_i (i.e. la séquence $\mathbf{x}_{a,\mathbf{t}_i}$) c'est le rendements y_{a,\mathbf{t}_i} qui suit. Ainsi, on veut entraîner un modèle capable de prédire avec une bonne prédiction la valeur de y_{a,\mathbf{t}_i} étant une séquence $\mathbf{x}_{a,\mathbf{t}_i}$ observé, par exemple le rendements d'un actif dans le mois qui suit une séquence de w mois.
- Par une méthode de rolling windows (fenêtre roulante), on prépare notre ensemble de périodes d'entraînement par l'équation

$$r = \{\mathbf{t}_i\}_{w \leq i \leq T_r} = \{\mathbf{t}_w, \mathbf{t}_{w+1}, \dots, \mathbf{t}_{T_r}\}, \quad (11)$$

où r désigne un tel ensemble⁴

²Dans le cas où nos points temporels désignent des mois, une période de taille w est simplement un ensemble de w mois.

³On peut aussi voir les séquences $\mathbf{x}_{a,\mathbf{t}_i}$ comme des vecteurs de dimensions w et on écrit $\mathbf{x}_{a,\mathbf{t}_i} \in^w$.

⁴La condition $w \leq i \leq T_r$ est là pour que le premier point temporel à utiliser soit $t_{w-w} = 0$ ($i = w$) et le dernier point temporel à utiliser soit t_{T_r-1} ($i = T_r$).

- Avec un ensemble de périodes d'entraînement r bien préparé, on construit notre ensemble d'inputs d'entraînement X_r à l'aide de l'équation

$$X_{a,r} = \{\mathbf{x}_{a,t}\}_{t \in r} = \{\mathbf{x}_{a,t_w}, \mathbf{x}_{a,t_w+1}, \dots, \mathbf{x}_{a,t_{T_r}}\}, \quad (12)$$

et ça représente un ensemble d'entrées (inputs) qu'on va utilisé pour entraîner nos modèles (kernels)⁵.

- On prépare nos outputs (étiquettes ou labels)—qu'on veut entraîner nos modèles de faire des prédictions qui leurs sont proches—par analogie à (12) avec l'équation

$$a,r = \{y_{a,t}\}_{t \in r} = \{y_{a,t_w}, y_{a,t_w+1}, \dots, y_{a,t_{T_r}}\} \quad (13)$$

- Par analogie, on utilise s pour désigner l'ensemble de périodes de test donné par

$$s = \{t_i\}_{t_{T_r} \leq i \leq t_T} = \{t_{T_r}, t_{T_r+1}, \dots, t_T\}. \quad (14)$$

Aussi par analogie aux équations (12) et (13), qu'on définit nos ensembles d'input de test $X_{a,s}$ et d'output de test a,s tel que

$$\begin{aligned} X_{a,s} &= \{\mathbf{x}_{a,t}\}_{t \in s} \\ a,s &= \{y_{a,t}\}_{t \in s}. \end{aligned} \quad (15)$$

- Notre ensemble d'entraînement par actif $a \in$ est alors $a,r = \{(\mathbf{x}_{a,t}, y_{a,t})\}_{t \in r}$ et celle de test est $a,s = \{(\mathbf{x}_{a,t}, y_{a,t})\}_{t \in s}$
- Finalement, notre ensemble d'entraînemenent pour tous les actifs de notre univers d'actifs comme l'union

$$\begin{aligned} r &= \bigcup_{a \in} a,r \\ &= \{(\mathbf{x}_{a,t}, y_{a,t})\}_{\substack{a \in \\ t \in r}} \end{aligned} \quad (16)$$

contenant tous les pairs input output d'entraînement possibles.

- Similairement, on définition notre ensemble de test pour tous les actifs de notre univers comme l'union

$$\begin{aligned} s &= \bigcup_{a \in} a,s \\ &= \{(\mathbf{x}_{a,t}, y_{a,t})\}_{\substack{a \in \\ t \in s}} \end{aligned} \quad (17)$$

3.3 Kernels comme mesure de similarité entre des séquences de rendements

- Un kernel $\kappa_\theta(\mathbf{x}, \mathbf{x}')$ où θ désigne les hyper-paramètres $\theta = \{\vartheta_1, \vartheta_2, \dots, \vartheta_{|\theta|}\}$ de κ ⁶, désigne une fonction prennaat deux inputs (deux séquences) et retournant un output, similarité entre $\mathbf{x} \in^w$ et $\mathbf{x}' \in^w$.

⁵L'entraînement d'un modèle est l'action d'optimiser, ou de trouver les meilleurs hyperparametres d'un modèle.

⁶ $|\theta|$ dénote combien d'éléments il y en a dans θ

- Par exemple le kernel Matérn ($\kappa_\theta^{(Ma)}$) a l'ensemble $\theta = \{\ell, \nu\}$, où ℓ désigne son length scale et ν le paramètre de smoothness (smoothing parameter), comme ensemble d'hyper-paramètres.
- On suppose qu'on dispose d'un ensemble de choix de kernels

$$\mathcal{K} = \left\{ \kappa_{\theta_1}^{(1)}, \kappa_{\theta_2}^{(2)}, \dots, \kappa_{\theta_m}^{(m)} \right\}. \quad (18)$$

Par exemple le cas où on travaille avec RBF, Matérn, Rational Quadratic, et Combined, on a $m = 4$, et chaque $\kappa_{\theta_i}^{(i)}$ désigne l'un de ces choix avec θ_i ses hyperparamètres.

- Plus spécifiquement, dans notre contexte un kernel $\kappa_\theta \in \mathcal{K}$ prend en entrée deux séquences de rendements d'un asset $a \in \mathbf{x}_{a,\mathbf{t}_i}$ et $\mathbf{x}_{a,\mathbf{t}_j}$ correspondant à deux périodes différentes \mathbf{t}_i et \mathbf{t}_j et il nous remet la similarité entre ces deux séquences de rendements:

$$(\mathbf{x}_{a,\mathbf{t}_i}, \mathbf{x}_{a,\mathbf{t}_j}) \mapsto \kappa_\theta(\mathbf{x}_{a,\mathbf{t}_i}, \mathbf{x}_{a,\mathbf{t}_j}). \quad (19)$$

3.4 Problème.

Le problème de la prédiction de rendement que nous considérons dans le cadre de ce projet est le suivant: étant donné un point temporel t dans l'ensemble de test, i.e $t \in_s$, un actif a tiré de l'univers d'actifs, décrire la densité de la distribution de probabilité postérieure du rendement $y_{a,t}$:

$$p(y_{a,\mathbf{t}} \mid \mathbf{x}_{a,\mathbf{t},a,r}), \quad (20)$$

avec $\mathbf{t} = \{t - w, \dots, t - 1\}$ ⁷.

4 Solution par la régression par processus gaussiens

4.1 Méthode de la régression par processus gaussiens

La régression par processus gaussiens (GPR, pour *Gaussian Process Regression*) est une méthode de régression probabiliste puissante, particulièrement adaptée aux problèmes où l'incertitude et la flexibilité du modèle sont essentielles. Elle permet non seulement de faire des prédictions moyennes, mais aussi de quantifier l'incertitude associée à ces prédictions.

4.1.1 Description globale de la méthodologie

La GPR repose sur l'hypothèse que les données peuvent être modélisées comme un processus gaussien, qui est une distribution de probabilités sur des fonctions. Mathématiquement, un processus gaussien est défini comme suit :

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \quad (21)$$

où $m(\mathbf{x})$ est la fonction moyenne (souvent supposée nulle pour simplifier), et $k(\mathbf{x}, \mathbf{x}')$ est la fonction de covariance ou noyau, qui encode les corrélations entre les points \mathbf{x} et \mathbf{x}' .

⁷La description d'une densité de probabilité se fait en générale en décrivant sa loi de probabilité.

Dans le cadre de la prédiction des rendements, l'objectif est d'estimer la probabilité conditionnelle suivante :

$$p(y_{a,t} \mid \mathbf{x}_{a,t,a,r}) \sim \mathcal{N}(\mu_{a,t}, \sigma_{a,t}^2), \quad (22)$$

où :

- $\mu_{a,t}$ est la prédiction moyenne, estimée comme la valeur attendue $\hat{\mathbb{E}}[R_{a,t}]$,
- $\sigma_{a,t}$ est l'écart-type ou volatilité associée à cette prédiction $\hat{\sigma}[R_{a,t}]$.

La GPR utilise les données d'entraînement $\mathbf{a}, r = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ pour définir une distribution a posteriori sur les fonctions. Cela implique de combiner une fonction moyenne $m(\mathbf{x})$ et un noyau $k(\mathbf{x}, \mathbf{x}')$ pour calculer les prédictions et leurs incertitudes.

4.1.2 Prédiction des rendements et de la volatilité

La GPR produit des prédictions sous forme de distribution gaussienne pour chaque point \mathbf{x}_t . Ces prédictions incluent une moyenne μ_t et une variance σ_t^2 , données par les formules suivantes :

1. Moyenne prédictive :

$$\mu_t = \mathbf{t}^\top (\mathbf{t}^\top + \sigma_n^2)^{-1}, \quad (23)$$

où :

- \mathbf{t} est le vecteur des covariances entre \mathbf{x}_t et les points d'entraînement,
- \mathbf{t}^\top est la matrice des covariances des points d'entraînement,
- σ_n^2 est la variance du bruit.

2. Variance prédictive :

$$\sigma_t^2 = k(\mathbf{x}_t, \mathbf{x}_t) - \mathbf{t}^\top (\mathbf{t}^\top + \sigma_n^2)^{-1} \mathbf{t}. \quad (24)$$

Ces deux quantités permettent de caractériser la distribution de probabilité des rendements $R_{a,t}$, avec $\mu_{a,t} = \hat{\mathbb{E}}[R_{a,t}]$ et $\sigma_{a,t} = \hat{\sigma}[R_{a,t}]$. La volatilité (écart-type) $\sigma_{a,t}$ est particulièrement utile pour évaluer le risque associé à la prédiction.

4.1.3 Applications aux rendements financiers

Dans le cadre de ce projet, la GPR est particulièrement utile car elle permet :

- de prédire les rendements futurs ($\mu_{a,t}$) tout en tenant compte des incertitudes ($\sigma_{a,t}$),
- d'analyser les risques associés aux investissements grâce aux intervalles de confiance,
- d'évaluer les relations complexes entre les actifs d'un portefeuille via des noyaux adaptés.

4.2 Optimisation des hyper-paramètres

L’optimisation des hyper-paramètres maximise la vraisemblance marginale logarithmique (LML) :

$$\log p(|, \theta) = -\frac{1}{2\theta} - \frac{1}{2} \log |\theta| - \frac{n}{2} \log 2\pi, \quad (25)$$

où θ est la matrice de covariance dépendant des hyper-paramètres θ .

L’algorithme **L-BFGS-B** optimizes the parameters of the noyau (e.g., length d’échelle, variance) by considering the contraintes, with $n = 40$ redémarrages to avoid the minima locaux, guaranteeing precision predictions and fiable uncertainties.

5 Expériences

5.1 Données

Les données utilisées dans ce projet ont été préparées comme suit :

5.1.1 Source des données

Les données financières ont été collectées à l’aide de la bibliothèque **yfinance**, un outil Python permettant d’extraire facilement les séries temporelles boursières. Cette bibliothèque fournit des données historiques, telles que les prix de clôture, les volumes, et autres indicateurs, directement depuis Yahoo Finance.

5.1.2 Choix de l’univers d’actifs

L’univers d’actifs est composé de grandes entreprises technologiques : IBM, META, AAPL, et MSFT. Ces choix sont justifiés par leur forte capitalisation boursière et leur représentativité dans des secteurs dynamiques, offrant ainsi une volatilité suffisante pour tester les modèles.

5.1.3 Ensemble temporel et fréquence

Les données couvrent la période de septembre 2019 à septembre 2022. Cette plage temporelle est choisie pour inclure des cycles de marché variés (pré-COVID, crise COVID, et reprise économique). La fréquence utilisée est quotidienne (*daily*), permettant de capturer les variations fines des rendements tout en garantissant une granularité suffisante pour l’entraînement des modèles.

5.2 Evaluation économique des résultats

Pour donner sens aux évaluations des performances prédictives discutées dans la section 2.3.1, on s’est demandé comment le choix du kernel influence le retour d’un portefeuille basé sur les prédictions fournis par nos kernels de choix. Un portefeuille π_t sur notre univers d’actifs est simplement un ensemble de poids qu’on décide à investir dans chaque actifs pendant la période $\{t - 1, t\}$ ⁸. Mathématiquement, un portefeuille π_t est décrit

⁸ $t - 1$ dénote le point temporel qui précède t .

par

$$\pi_t = \{w_{a,t}\}_{a \in \mathcal{A}} = \{w_{a_1,t}, \dots, w_{a_n,t}\}, \quad (26)$$

où chaque $w_{a,t}$ décrit le pourcentage de notre budget on investit dans un actif $a \in \mathcal{A}$ durant la période $\{t-1, t\}$. On met la condition $\sum_{a \in \mathcal{A}} w_{a,t} = 1$ afin de ne pas excéder notre budget B (B pouvant être n'importe quel montant).

La façon avec laquelle on définit les valeurs de $w_{a,t}$ est ce qu'on appelle *construction de portefeuille*. Trouver la méthode optimale est une question ouverte de recherche [1], et sort du cadre du projet. Dans notre cas on a fixé notre stratégie de construction de portefeuille suivant un schéma de pondération heuristique, inspiré du ratio de Sharpe [2] et des concepts de parité des risques, où les pondérations des actifs individuels sont allouées proportionnellement à leurs rendements ajustés au risque attendus [3].

Mathématiquement, étant la prédiction du rendement attendu $\mu_{a,t}$ et le risque attendu $\sigma_{a,t}$ pour un actif $a \in \mathcal{A}$, notre pondération heuristique commence d'abord par calculer le score

$$s_{a,t} = \frac{\mu_{a,t}}{\sigma_{a,t}}, \quad (27)$$

qui est une division du rendement attendu de l'actif a sur son risque attendu. Plus notre rendement attendu est grand et notre risque est petit, plus notre score $s_{a,t}$ sera grand. À l'aide d'un tel score, on procède à la génération des poids à investir durant la période $t-1$ et t comme

$$w_{a,t} = \frac{s_{a,t}}{\sum_{a' \in \mathcal{A}} s_{a',t}} \quad (28)$$

pour tout actif $a \in \mathcal{A}$. La normalisation en divisant par $\sum_{a' \in \mathcal{A}} s_{a',t}$ assure que la condition $\sum_{a \in \mathcal{A}} w_{a,t} = 1$.

Ainsi, on peut évaluer la performance économique (retour) de nos m portefeuilles construits à partir des prédictions générés par nos différents kernels $\kappa \in \mathcal{K}$ pour toutes les points temporels de test $t_j \in \mathcal{T}_s$. Ceci veut dire que pour chaque point futur $t_j \in \mathcal{T}_s$, on prédit le retour attendu et le risque à ce point et on forme le portefeuille

$$\pi_t^{(\kappa)} = \left\{ w_{a,t}^{(\kappa)} \right\}_{a \in \mathcal{A}}, \quad (29)$$

tel que les poids sont calculés à l'aide des prédictions d'un kernel $\kappa \in \mathcal{K}$, qu'on investit entre $t-1$ et t .

Finalement, pour chaque kernel $\kappa \in \mathcal{K}$, on aura une séquence de rendements *réalisés* (rendements après l'effet) de portefeuilles à chaque instant $t \in \mathcal{T}_s$ calculé par

$$R_{\pi,t}^{(\kappa)} = \sum_{a \in \mathcal{A}} w_{a,t}^{(\kappa)} \cdot R_{a,t}, \quad (30)$$

où $R_{a,t}$ dénotent les rendements réalisés calculé par (10). On dénote la séquence de rendements réalisés par un portefeuille $\pi^{(k)}$ par

$$R_{\pi,s}^{(\kappa)} = \left(R_{\pi,t}^{(\kappa)} \right)_{t \in \mathcal{T}_s}, \quad (31)$$

et on calcule le rendement du portefeuille $\pi^{(\kappa)}$ durant toute la période s donné par la relation

$$(1 + R_{\pi,s}^{(\kappa)}) = \prod_{t \in \mathcal{T}_s} (1 + R_{\pi,t}^{(\kappa)}), \quad (32)$$

où $R_{\pi,s}^{(\kappa)}$ dénote un tel rendement [4].

La métrique usuelle avec laquelle on mesure la performance économique d'un portefeuille durant la période d'investissement s est le rendement de Sharpe [2]. Celle-ci est calculée comme la division du retour du portefeuille durant cette période divisée par le risque (volatilité) réalisée du portefeuille durant une telle période. Mathématiquement ceci se traduit par

$$S_{\pi,s}^{(\kappa)} = \frac{R_{\pi,s}^{(\kappa)}}{\sigma \left[\begin{smallmatrix} (\kappa) \\ \pi,s \end{smallmatrix} \right]}, \quad (33)$$

où $\sigma \left[\begin{smallmatrix} (\kappa) \\ \pi,s \end{smallmatrix} \right]$ désigne le risque réalisé du portefeuille $\pi^{(\kappa)}$ durant la période s . Plus la métrique $S_{\pi,s}^{(\kappa)}$ est élevée le mieux notre portefeuille $\pi^{(\kappa)}$ a performé durant la période s puisque c'est celui qui coïncide avec le portefeuille qui a réaliser un bon rendement et moins de risque.

5.3 Analyse des résultats expérimentaux

Le tableau 2 compare les performances des noyaux $\kappa \in \mathcal{K}$ en termes de rendement total et de ratio de Sharpe pour le portefeuille multi-actifs (IBM, MSFT, ADBE, PYPL, CRM).

| Kernel | Rendement Total | Ratio de Sharpe |
|--------------------|-----------------|-----------------|
| RBF | 13.96% | 0.3402 |
| Matérn | 10.84% | 0.3028 |
| Rational Quadratic | 11.59% | 0.3091 |
| Combined | 17.21% | 0.3823 |

Table 2: Performance des noyaux sur un portefeuille multi-actifs. Les meilleures performances sont en gras.

Les principaux résultats sont décrits comme suit :

- Le noyau **Combined** se distingue avec le rendement total le plus élevé (**17.21%**) et le meilleur ratio de Sharpe (**0.3823**), démontrant une excellente capacité à équilibrer rendement et risque.
- Le noyau **RBF** montre un rendement compétitif (13.96%) mais un ratio de Sharpe inférieur (0.3402), indiquant une volatilité plus élevée par rapport aux gains réalisés.
- Les noyaux **Matérn** et **Rational Quadratic** ont des performances similaires, avec des ratios de Sharpe modérés (0.3028 et 0.3091), ce qui suggère une capacité moyenne à capturer les relations inter-actifs.
- Globalement, les résultats confirment que la combinaison des noyaux (**Combined**) améliore la robustesse et capture efficacement les dépendances temporelles et inter-actifs.

6 Conclusion

Ce travail a démontré l'efficacité des processus gaussiens pour la prédiction des rendements financiers, en évaluant différents noyaux (RBF, Matérn, Rational Quadratic et Combined). Les résultats expérimentaux montrent que le noyau **Combined** surpasse les autres en termes de rendement total (**17.21%**) et de ratio de Sharpe (**0.3823**), attestant de sa capacité à capturer les dépendances complexes entre les actifs et les variations temporelles.

En plus des prédictions, l'approche GPR offre un avantage unique : elle quantifie l'incertitude à travers des intervalles de confiance, permettant ainsi une gestion des risques plus rigoureuse. Cette capacité est cruciale dans un domaine où la précision seule ne suffit pas.

Ce travail constitue une **baseline** pour la recherche dans ce domaine pour plusieurs raisons :

- Il établit une méthodologie robuste combinant des noyaux adaptés et une optimisation fiable des hyper-paramètres.
- Les résultats fournissent des benchmarks clairs pour évaluer d'autres modèles, qu'ils soient basés sur des approches non probabilistes ou deep learning.
- L'étude met en lumière l'importance des noyaux dans la capture des relations interactifs et des dynamiques temporelles, ouvrant la voie à des recherches sur des noyaux plus sophistiqués ou spécifiques.

En conclusion, cette recherche pose les bases pour des études futures visant à améliorer la prédiction des rendements financiers, en intégrant des approches hybrides (modèles probabilistes et réseaux neuronaux) ou en étendant l'analyse à des portefeuilles plus diversifiés et dynamiques.

References

- [1] Bryan Kelly, Dacheng Xiu, et al. Financial machine learning. *Foundations and Trends® in Finance*, 13(3-4):205–363, 2023.
- [2] William F Sharpe. Mutual fund performance. *The Journal of business*, 39(1):119–138, 1966.
- [3] Clifford S Asness, Andrea Frazzini, and Lasse H Pedersen. Leverage aversion and risk parity. *Financial Analysts Journal*, 68(1):47–59, 2012.
- [4] Eric Jondeau, Ser-Huang Poon, and Michael Rockinger. *Financial modeling under non-Gaussian distributions*. Springer Science & Business Media, 2007.