Université Paris-Dauphine.

Master 2. Research in Finance

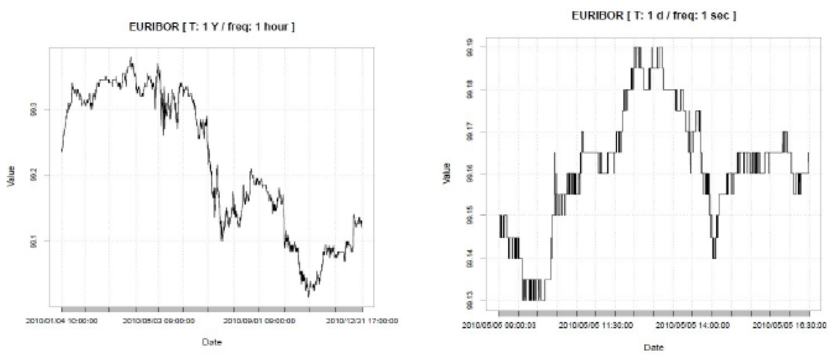
Projet. Microstructure des prix financiers.

Irina Kortchemski

Introduction. Carnet d'ordre limite.

Les propriétés des trajectoires du prix d'un actif coté sur un marché boursier dépendent fortement de la fréquence d'observation de la trajectoire. Si pour des observations à basse fréquences (échelle de temps : jour, mois) les trajectoires sont bien approchées par des processus de diffusion (par example, le mouvement logonormale, basé sur le mouvement Brownien)

Pour des fréquences plus élevées (échelle de temps : heure, minute, ou en dessous), le prix a une nature discrète et est représenté par des trajectoires discontinues, constantes par morceaux qui varient sur une grille. L'amplitude minimale de variation du prix à chaque sauts (0 :01 dans la Figure 1) est appelée tick.



(a) (b)

Figure 1. Trajectoires du prix de l'Euribor future pour différentes fréquences d'observation. Pham (2013), présentation au NUS-UTokyo Workshop on Quantitative Finance.

Suivant une approche typique dans la modélisation, on décrit ce prix qui varie sur grille discrète par le biais d'une processus ponctuel $(T_n, J_n)_{n\geq 1}$. La suite croissante $(T_n)_n$ représente les instants de variations (sauts d'un ou plusieurs ticks) du prix. Les v.a. J_n prennent leurs valeurs dans l'ensemble fini

$$E = \{+1, -1, ..., +m, -m\} \subset \mathbb{Z}\backslash\{0\}. \tag{1}$$

et représentent les incréments (positifs ou négatifs) du prix. Le prix de l'actif est donné par le processus (constant par morceaux).

$$P_t = P_0 + \sum_{k=1}^{N_t} J_k \tag{2}$$

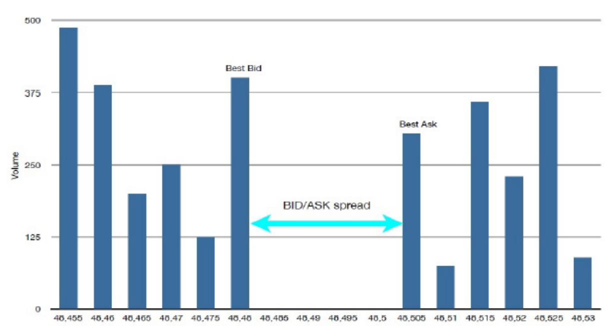


Figure 2. Instantané d'un carnet d'ordre limite (limit order book).

où N_t est le nombre d'occurrences avant t,

$$N_t = \sum_{n\geq 1} 1_{\{T_n\leq t\}}.$$

La valeur courante P_t représente la dernière cotation du prix mid, soit la moyenne entre le meilleur ask et le meilleur bid, Figure 4.

2. Une modélisation simplifiée.

Dans le cadre d'une première modélisation simplifiée, on suppose que la suite des incréments $(J_n)_n$ est i.i.d. (absence de dépendance entre les variations du prix). On suppose de plus que les instants de sauts (T_n) sont indépendants des $(J_n)_n$, et que la suite des durées des intervalles entre sauts

$$S_n = T_{n+1} - T_n$$

est i.i.d. de loi exponentielle de paramètre λ.

Q1. On souhaite estimer la probabilité que dans le modèle ainsi défini les trajectoires du prix puissent prendre des valeurs négatives,

$$\mathbb{P}(\inf_{t\leq T} P_t < 0)$$

pour $P_0 = 10$ et $T = 4h$. On prendra une valeur de λ telle que la durée moyenne entre un saut et le suivant soit autour de 300s. On considérera les cas m = 1 et m = 3 avec loi des sauts symétrique : $\mathbb{P}(J_n = 1) = \mathbb{P}(J_n = -1) = 1/2$ pour $m = 1$ et

$\mathbb{P}(J_n = \pm i) = p_i$ pour $i \in E$ et $m = 3$, où $p_1 = 1/4, p_2 = 1/6, p_3 = 1/12$.

3. Prise en compte d'autres propriétés des données haute fréquence.

Il est important qu'un modèle de prix d'actifs à haute fréquence soit capable de reproduire, en plus de la nature discrète des prix, les faits stylisés suivants :

- Retour vers la moyenne : les variations du prix sur des courtes échelles de temps sont fortement anti-corrélées. Cela détermine un effet appelé retour vers la moyenne : le prix, tout en étant soumis à des variations aléatoires, a tendance à osciller autour d'une valeur intermédiare. Ce phénomène est en lien avec la structure du carnet d'ordre : l'exécution d'un ordre à l'achat ou à la vente consomme l'un de deux cotés du carnet. Si tout le volume à disposition au meilleur prix bid ou au meilleur prix ask est entièrement epuisé par l'exécution de l'ordre, l'effet est d'augmenter le best ask, ou de diminuer le best bid (et par conséquent le prix mid change aussi).
- Clusters de volatilité : on observe une alternance de périodes avec activité de trading intense (volatilité du prix plus élévée) et de périodes plus calmes (volatilité plus faible). Par conséquent, les instants de sauts ne sont pas distribués de manière uniforme.

1. Afin de rajouter une plus forte composante de retour vers la moyenne, on peut modéliser le prix mid par une superposition de processus

$$P_t = P_t^{(1)} + P_t^{(2)},$$

où

$$P_t^{(i)} = P_0/2 + \sum_{k=1}^{N_t^{(i)}} J_k^{(i)} \quad i = 1, 2$$

Toutes les variables aléatoires indexées par $i = 1$ et $i = 2$ étant indépendantes.

(1) : on fera les mêmes hypothèses que dans la section précédente pour les distributions de ${N_t^{(1)}}_{t\geq 0}$ et ${N_t^{(2)}}_{t\geq 0}$, en changeant la durée moyenne $1/\lambda_1$ entre les sauts de 300s à 660s.

(2) : la suite de sauts $J_{n+1}^{(2)} = -J_n^{(2)}$, à partir de $J_1^{(2)}$ avec

$$\mathbb{P}(J_1^{(2)} = 1) = \mathbb{P}(J_1^{(2)} = 1) = 1/2,$$

modélisera la composante soumise à oscillation. On prendra pour ${N_t^{(2)}}_{t\geq 0}$ un processus de Poisson simple, de vie moyenne entre sauts $1/\lambda_2 = 110$s.

Q2. On veut reprendre les points de la question précédente afin d'estimer l'impacte d'une telle modélisation.

II. Une autre approche à la modélisation du phenomène de retour vers la moyenne consiste à s'affranchir de l'hypothèse d'indépendance entre les sauts. On décompose à ce fin la variation du prix à chaque saut en

$$J_n = \widehat{J}_n \xi_n$$

où

$$\widehat{J}_n = sign(J_n) \quad et \quad \xi_n = |J_n|.$$

On supposera maintenant que les variations du prix ne sont pas indépendantes entre elles, mais ce n'est que le signe d'un incrément de prix (et pas sa valeur absolue) qui influence les incréments futures, de plus, le signe de l'incrément à l'instantT_n ne dépend que de celui en T_{n-1}, et pas des précédents. Mathématiquement, cela se traduit en prenant pour $\widehat{J}_n$ une chaîne de Markov, de matrice de transition

où I_n est une chaîne de Markov de matrice de transition

$$\hat{Q} = \begin{pmatrix} \frac{1+\alpha_+}{2} & \frac{1-\alpha_+}{2} \\ \frac{1-\alpha_-}{2} & \frac{1+\alpha_-}{2} \end{pmatrix} \tag{3}$$

avec $\alpha_+, \alpha_- \in [-1, 1)$.

Les $(\xi_n)_n$ sont i.i.d. à valeurs dans $\{1, 2, ...m\}$ indépendants des $(\hat{J}_n)_n$, de loi

$$\mathbb{P}(\xi_n = i) = p_i, \quad i = 1, 2, ...m$$

On modélise de cette manière des amplitudes de saut du prix qui sont indépendantes entre elles, et indépendantes de la direction du saut (hausse ou baisse). Cette propriété traduit une indépendance des volumes échangés dans le carnet d'ordre à chaque événement de trading, et leur indépendance par rapport au type de transaction (vente ou achat).

Q. Quel signe faudra-t-il choisir pour les paramètres $\alpha_\pm$ afin de modéliser le phénomène de retour vers la moyenne ?

On gardera par ailleurs la modélisation (4) pour le prix, avec l'hypothèse d'indépendance entre $(J_n)_n$ et $(T_n)_n$, et une distribution exponentielle pour les $(S_n)_n$.

Q3. Des estimations à partir des données sur l'Euribor future [FP] On fixera pour $\alpha = -0.875$ la valeur estimée dans l'article [FP].

Simuler des trajectoires du prix P_t, observer empiriquement sur les simulations que l'on a

$$\mathbb{P}[\hat{J}_n \cdot \hat{J}_{n+1}] \sim 0.1.$$

Estimer dans ce cas également la probabilité que le prix P_t prenne des valeurs négatives au cours de $N = 100$ variations.

4. Limite macroscopique .

On souhaite étudier le comportement des prix engendré par le modèle considéré dans la question précédente sur des échelles de temps plus longues. En supposant une durée de 8h pour chaque jour et 250 jours de trading par an, horizon $T = 1$ an, estimer la probabilité que le prix P_t prenne des valeurs négatives, ainsi que les quantiles à 10^{-6} de la valeur finale du prix, à partir de la valeur $P_0 = 1000$ (typique pour un indice boursier) et à un horizon $T = 1$ an.

5. Réferences.

Pietro FODRA, Huyen PHAM. "Semi Markov model for market microstructure".