Données haute fréquence s

Analyse et modélisation statistique multi-échelle de séries chronologiques financières

Cours de Master - Probabilités et Finances - Sorbonne Université'

Slides de la partie III

Tick by tick financial time series - Digression sur les Processus de Hawkes

Emmanuel Bacry
DR CNRS, CEREMADE, Université Paris-Dauphine, PSL,
CSO, Health Data Hub
emmanuel.bacry@cnrs.fr

- Point processes introduced by A.G.Hawkes in the 70's
- Flexible and versatile tool to investigate mutual and/or self interaction of dynamic flows
- Very sucessful in seismic (> 1980)
- Rising popularity in finance (> 2007)
  $\longrightarrow$ Modeling high frequency time-series events (price changes, cancel/limit/market orders, $\ldots$)
- Rising popularity in machine learning (network, ...)

The 1-Dimensional Poisson process

- N_t : a jump process (jumps are all of size 1)
- λ_t : intensity ($\simeq$ density of jumps)

$$\lambda_{\mathbf{t}} = \mu$$

$\Longrightarrow$ **The inter-arrival times are independant**

The Poisson process

- N_t : jump process (jumps are all of size 1)
- λ_t : the intensity
- μ : 1-dimensional exogenous intensity

$$\lambda_\mathbf{t} = \mu$$

A Hawkes process

$\Longrightarrow$ Introducing (positive) correlation in the arrival flow

$\Longrightarrow$ "Auto-regressive" relation

$$\lambda_\mathbf{t} = \mu + \phi \star \mathbf{dN_t},$$

where by definition

$$\phi \star dN_t = \int_{-\infty}^{+\infty} \phi(t-s)dN(s)$$

and $\phi(t)$: kernel function, positive and causal (supported by R^+).

The Poisson process

- N_t : jump process (jumps are all of size 1)
- λ_t : the intensity
- μ : 1-dimensional exogenous intensity

$$\lambda_\mathbf{t} = \mu$$

A Hawkes process

$\Longrightarrow$ Introducing (positive) correlation in the arrival flow
$\Longrightarrow$ "Auto-regressive" relation

$$\lambda_\mathbf{t} = \mu + \phi \star \mathbf{dN_t},$$

where by definition

$$\phi \star dN_t = \int_{-\infty}^{+\infty} \phi(t-s) dN(s)$$

and $\phi(t)$: kernel function, positive and causal (supported by R^+).

Hawkes processes - general definition in dimension D

- N_t : a D-dimensional jump process (jumps are all of size 1)
- λ_t : D-dimensional stochastic intensity
- μ : D-dimensional exogenous intensity
- $\Phi(t)$: $D \times D$ square matrix of kernel functions $\Phi^{ij}(t)$ which are positive and causal (i.e., supported by R^+).

"Auto-regressive" relation

$$\lambda_\mathbf{t} = \mu + \mathbf{\Phi} \star \mathbf{dN_t},$$

where by definition

$$(\Phi \star dN_t)^i = \sum_{k=1}^{D} \int_{-\infty}^{+\infty} \Phi^{ik}(t-s) dN^k(s)$$

Stationnarity of λ_t ? ($\mathbb{E}(\lambda_t)$ should not depend on t)

Since

$$\lambda_{\mathbf{t}} = \mu + \mathbf{\Phi} \star \mathbf{dN_t},$$

One gets

$$\mathbb{E}(\lambda_t) = \mu + \Phi \star \mathbb{E}(\lambda_t)$$

which implies

$$(\delta_t \mathbb{I} - \Phi_t) \star \mathbb{E}(\lambda_t) = \mu$$

We take the Fourier transform ($\hat{f}(i\omega) = \int f(t) e^{-i\omega t} dt$)

$$(\mathbb{I} - \hat{\Phi}(i\omega))\widehat{\mathbb{E}(\lambda_t)}(i\omega) = 2\pi\delta(\omega)\mu$$

Stationnarity of λ_t ?

First step : **Find a condition for $\mathbb{E}(\lambda_t)$ not to depend on t ?**

$$(\mathbb{I} - \hat{\Phi}(i\omega))\widehat{\mathbb{E}(\lambda_t)}(i\omega) = 2\pi\mu\delta(\omega)$$

If the operator $(\mathbb{I} - \hat{\Phi}(i\omega))$ was inversible for all ω we could write

$$\widehat{\mathbb{E}(\lambda_t)}(i\omega) = 2\pi\delta(\omega)(\mathbb{I} - \hat{\Phi}(i\omega))^{-1}\mu$$

which implies that $\widehat{\mathbb{E}(\lambda_t)}(i\omega)$ is proportional to a dirac function

which implies that $\mathbb{E}(\lambda_t)$ is a constant function

Proposition

$$\rho(\hat{\Phi}(0)) < 1 \Longrightarrow (\mathbb{I} - \hat{\Phi}(i\omega)) \text{ is invertible } \forall\omega$$

where $\rho(M)$ is the spectral radius of the matrix M (i.e., the maximum eigenvalue modulus), i.e.,

$$\rho(M) = \sup_x \frac{||Mx||_2}{||x||_2}$$

Proof 1/2

Let ω such that $(\mathbb{I} - \hat{\Phi}(i\omega))$ is not invertible, i.e.,
$\exists x^*, \quad \hat{\Phi}(i\omega)x^* = x^*$
$\Rightarrow \forall k, \quad \sum_l \hat{\Phi}^{kl}(i\omega)x_l^* = x_k^*$
$\Rightarrow \forall k, \quad \sum_l |\hat{\Phi}kl(i\omega)||x_l^*| \geq |x_k^*|$

Since $\Phi^{kl}(t) \geq 0$,

$$
\begin{aligned}
|\hat{\Phi}^{kl}(i\omega)| &= |\int \Phi^{kl}(t)e^{-i\omega}dt| \\
&\leq \int |\Phi^{kl}(t)|dt \\
&= \int \Phi^{kl}(t)dt = \hat{\Phi}^{kl}(0)
\end{aligned}
$$

Thus $\forall k, \quad \sum_l \hat{\Phi}^{kl}(0)|x_k^*| \geq |x_l^*|$,

Proof 2/2

We proved $\forall k, \ \ \sum_l \hat{\Phi}^{kl}(0)|x_k^*| \geq |x_l^*|$,

Thus

$\forall k, \ \ |\sum_l \hat{\Phi}^{kl}(0)|x_k^*||^2 \geq |x_l^*|^2$,

$\Rightarrow \sum_k |\sum_l \hat{\Phi}^{kl}(0)|x_l^*||^2 \geq \sum_k |x_k^*|^2$,

$\Rightarrow \exists x, \ ||\hat{\Phi}(0)x||_2 \geq ||x||^2$,

$\Rightarrow \exists x, \frac{||\hat{\Phi}(0)x||_2}{||x||_2} \geq 1$

$\Rightarrow \rho(\hat{\Phi}(0)) \geq 1$

So we proved the proposition, i.e.,

$\exists \omega$ such that $(\mathbb{I} - \hat{\Phi}(i\omega))$ is not invertible, $\Rightarrow \rho(\hat{\Phi}(0)) \geq 1$

Thus we proved (Proposition)

$$\rho(\hat{\Phi}(0)) < 1 \Longrightarrow (\mathbb{I} - \hat{\Phi}(i\omega)) \text{ is invertible } \forall\omega$$

Remember that we also proved

$$(\mathbb{I} - \hat{\Phi}(i\omega))\widehat{\mathbb{E}(\lambda_t)}(i\omega) = 2\pi\delta(\omega)\mu$$

Thus

$$\rho(\hat{\Phi}(0)) < 1 \Longrightarrow \widehat{\mathbb{E}(\lambda_t)}(i\omega) = 2\pi\delta(\omega)(\mathbb{I} - \hat{\Phi}(0))^{-1}\mu$$

Thus if $\rho(\hat{\Phi}(0)) < 1$ then $\mathbb{E}(\lambda_t)$ is the constant function

$$\Lambda = \mathbb{E}(\lambda_t) = (\delta\mathbb{I} - \hat{\Phi}(0))^{-1}\mu$$

Hawkes processes - Stationnarity

Theorem (not proved here)

If $\rho(\hat{\Phi}(0)) < 1$ then λ_t is a stationnary process

Morever

$$\Lambda = \mathbb{E}(\lambda_t) = (\delta\mathbb{I} - \hat{\Phi}(0))^{-1}\mu$$

Let us introduce a useful notation : we set $\Psi(t)$ the matrix such that

$$\mathbb{I} + \hat{\Psi}(i\omega) = (\mathbb{I} - \hat{\Phi}(i\omega))^{-1},$$

i.e., $\hat{\Psi}(i\omega) = \hat{\Phi}(i\omega) + \hat{\Phi}(i\omega)^2 + \hat{\Phi}(i\omega)^3 + \ldots$

Consequently

- $\Psi(t) = \Phi(t) + \Phi(t) \star \Phi(t) + \Phi(t) \star \Phi(t) \star \Phi(t) + \ldots$
- $\Psi(t)$ is a causal function
- $\Psi(t)$ and $\Phi(t)$ commute
- If $\rho(\hat{\Phi}(0)) < 1$ then

$$\Lambda = \mathbb{E}(\lambda_t) = (\delta\mathbb{I} + \hat{\Psi}(0))\mu$$

A martingale representation

We define the infinitesimal martingale dM_t as :

$$dM_t = dN_t - \lambda_t dt$$

Thus

$$\lambda_t = \mu + \Phi \star dN_t = \mu + \Phi \star dM_t + \Phi \star \lambda_t$$

Consequently

$$(\delta \mathbb{I} - \Phi) \star \lambda_t = \mu + \Phi \star dM_t$$

which gives (when $\rho(\hat{\Phi}(0)) < 1$)

$$\boxed{\lambda_t = \Lambda + \Psi \star dM_t}$$

In the case the process is stationnary ($\rho(\hat{\Phi}(0)) < 1$), we have proved

$$\lambda_t = \Lambda + \Psi \star dM_t$$

Now, we would like to characterize second order statistics. For that purpose we propose to cpmpute the **infinitesimal covariance** defined as

$$C(t' - t)dtdt' = \mathbb{E}(dN_t dN_{t'}^T) - \Lambda\Lambda^T dtdt'$$

It contains all the second-order statistics information of the Hawkes process.

Computing the infinitesimal covariance

$$C(t'-t)dtdt' = \mathbb{E}(dN_t dN_{t'}^T) - \Lambda\Lambda^T dtdt'$$

using the martingale dM_t defined as

$$dN_t = dM_t + \lambda_t dt$$

we get 4 terms

$$\begin{aligned} C(t'-t)dtdt' &= \mathbb{E}((dM_t + \lambda_t dt)(dM_{t'}^T + \lambda_{t'}^T dt')) - \Lambda\Lambda^T dtdt' \\ &= \mathbb{E}(dM_t dM_{t'}^T) + \mathbb{E}(\lambda_t dM_{t'}^T)dt + \mathbb{E}(dM_t \lambda_{t'}^T)dt' + \mathbb{E}(\lambda_t \end{aligned}$$

First term $\mathbb{E}(dM_t dM_{t'}^T)$

- if $t \neq t'$, $\mathbb{E}(dM_t dM_{t'}^T) = 0$
- if $t = t'$ and $i \neq j$, $\mathbb{E}(dM_t^i dM_t^j) = 0$
- if $t = t'$ and $i = j$, $\mathbb{E}(dM_t^i dM_t^i) = \mathbb{E}(dN_t^i dN_t^i) = \Lambda^i dt$

Thus

$$\mathbb{E}(dM_t dM_{t'}^T) = \Sigma\delta(t' - t)dt$$

where $\delta(t)$ is the dirac distribution and

$$\Sigma = \begin{pmatrix} \Lambda^1 & 0 & \cdots & 0 \\ 0 & \Lambda^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Lambda^D \end{pmatrix}$$

Second term $\mathbb{E}(\lambda_t dM_{t'}^T)dt$
Using $\lambda_t = \Lambda + \Psi \star dM_t$, one gets

$$\begin{aligned}\mathbb{E}(\lambda_t dM_{t'}^T)dt &= \mathbb{E}(\Psi \star dM_t dM_{t'}^T)dt \\ &= dt \int_{s<t} \Psi(t-s)\mathbb{E}(dM_s dM_{t'}^T)\end{aligned}$$

Since we just proved that $\mathbb{E}(dM_t dM_{t'}^T) = \Sigma\delta(t'-t)dt'$

$$\begin{aligned}\mathbb{E}(\lambda_t dM_{t'}^T)dt &= dtdt' \int_{s<t} \Psi(t-s)\Sigma\delta(t'-s) \\ &= \Psi(t-t')\Sigma dtdt'\end{aligned}$$

Third term $\mathbb{E}(dM_t \lambda_{t'}^T)dt'$
Using second term computation $\mathbb{E}(\lambda_t dM_{t'}^T)dt = \Psi(t-t')\Sigma dt dt'$, one gets

$$\begin{aligned}
\mathbb{E}(dM_t \lambda_{t'}^T)dt' &= \mathbb{E}(\lambda_{t'} dM_t^T)^T dt' \\
&= (\Psi(t'-t)\Sigma)^T dt dt' \\
&= \Sigma \Psi(t'-t)^T dt dt'
\end{aligned}$$

Fourth term $\mathbb{E}(\lambda_t \lambda_{t'}^T) - \Lambda\Lambda^T$

Using $\lambda_t = \Lambda + \Psi \star dM_t$, one gets

$$\begin{aligned}\mathbb{E}(\lambda_t \lambda_{t'}^T) - \Lambda\Lambda^T &= \mathbb{E}\left(\int \Psi(t-s)dM_s \int dM_u^T \Psi(t'-u)^T\right) \\ &= \int\int \Psi(t-s)\mathbb{E}(dM_s dM_u^T)\Psi(t'-u)^T)\end{aligned}$$

Since we just proved that $\mathbb{E}(dM_t dM_{t'}^T) = \Sigma\delta(t'-t)dt'$

$$\begin{aligned}\mathbb{E}(\lambda_t \lambda_{t'}^T) - \Lambda\Lambda^T &= \int\int \Psi(t-s)\Sigma\delta(s-u)ds\Psi(t'-u)^T \\ &= \int \Psi(t-s)\Sigma\Psi(t'-s)^T ds \\ &= \int \Psi(t-s)\Sigma\Psi(t'-s)^T ds \\ &= \tilde{\Psi} \star \Sigma\Psi^T(t'-t)d\tau\end{aligned}$$

where $\tilde{\Psi}(s) = \Psi(-s)$

Finally

$$\begin{aligned} C(t'-t)dtdt' &= \mathbb{E}(dM_t dM_{t'}^T) + \mathbb{E}(\lambda_t dM_{t'}^T)dt + \mathbb{E}(dM_t \lambda_{t'}^T)dt' + \mathbb{E}(\lambda_t \\ &= \Sigma\delta(t'-t)dt + \Psi(t-t')\Sigma dtdt' + \Sigma\Psi(t'-t)^T dtdt' \end{aligned}$$

Thus

$$\boxed{C(\tau) = (\delta\mathbb{I} + \tilde{\Psi}) \star \Sigma(\delta\mathbb{I} + \Psi^T)(\tau)}$$

with the convention $\delta \star \delta = \delta$

We get an explicit formula for the second-order statistics in the stationnary case !

Hawkes processes : Estimation

We just proved, in the case λ_t is stationnary that the second-order statistics can be obtained through

$$C(\tau) = (\delta\mathbb{I} + \tilde{\Psi}) \star \Sigma(\delta\mathbb{I} + \Psi^T)(\tau)$$

Important question : **Could we use this equation for estimation of the kernel functions Φ ?**

Idea : Applying the Fourier transform on

$$C(\tau) = (\delta\mathbb{I} + \tilde{\Psi}) \star \Sigma(\delta\mathbb{I} + \Psi^T)(\tau)$$

one gets (using the Laplace transform $\hat{f}(z) = \int f(t)e^{-zt}dt$)

$$\hat{C}(z) = (\mathbb{I} + \hat{\Psi}(-z))\Sigma(\mathbb{I} + \hat{\Psi}^T(z))$$

Particular case : $D = 1$ (1-dimensional Hawkes process), one gets

$$\hat{C}(i\omega) = (\mathbb{I} + \hat{\psi}^*(i\omega))\Lambda(\mathbb{I} + \hat{\psi}(i\omega)) = \Lambda|\mathbb{I} + \hat{\psi}^*(i\omega)|^2$$

Pb : Given $\hat{C}$ is there a unique function $\hat{\psi}$ statisfying this last equation ? s

The quadratic form makes us think there are multiple solutions ... ?

A nice trick . . . :

We (right) multiply the equation

$$\hat{C}(z) = (\mathbb{I} + \hat{\Psi}(-z))\Sigma(\mathbb{I} + \hat{\Psi}^T(z))$$

on each side by $(\mathbb{I} + \hat{\Psi}^T(z))^{-1} = \mathbb{I} - \hat{\Phi}^T(z)$, and we get

$$\hat{C}(z)(\mathbb{I} - \hat{\Phi}^T(z)) = (\mathbb{I} + \hat{\Psi}(-z))\Sigma$$

which is equivalent to

$$C \star (\mathbb{I}\delta - \Phi^T)(\tau) = (\mathbb{I}\delta(\tau) + \tilde{\Psi}(\tau))\Sigma$$

Since $\Psi(\tau)$ is causal, $\tilde{\Psi}(\tau)$ is anti-causal, and consequently

$$C \star (\mathbb{I}\delta - \Phi^T)(\tau) = 0, \quad \text{for} \quad \tau > 0$$

From

$$C \star (\mathbb{I}\delta - \Phi^T)(\tau) = 0, \quad \text{for} \quad \tau > 0$$

We finally get a linear equation !

$$C(\tau) = C \star \Phi^T(\tau), \quad \text{for} \quad \tau > 0$$

Remember C has a dirac component ($C(\tau) = (\delta\mathbb{I} + \tilde{\Psi}) \star \Sigma(\delta\mathbb{I} + \Psi^T)(\tau)$), we can set $g(\tau)$ a function such that

$$g(\tau) = C^T(\tau)\Sigma^{-1} - \mathbb{I}\delta(\tau)$$

we finally get

$$\Sigma g^T(\tau) = (\Sigma g^T(\tau) + \Sigma\delta(\tau)) \star \Phi^T(\tau), \quad \text{for} \quad \tau > 0$$

and

$$\boxed{g(\tau) = \Phi(\tau) + \Phi \star g(\tau), \quad \text{for} \quad \tau > 0}$$

Recap

- The infinitesimal covariance is defined as

$$C(t'-t)dtdt' = \mathbb{E}(dN_t dN_{t'}^T) - \Lambda\Lambda^T dtdt'$$

- We setted $g(\tau) = C^T(\tau)\Sigma^{-1} - \mathbb{I}\delta(\tau)$
- we finally got

$$\boxed{g(\tau) = \Phi(\tau) + \Phi \star g(\tau), \quad \text{for} \quad \tau > 0}$$

Estimation Pb : We can compute $g(\tau)$

- How do we get from $g(\tau)$ to $\Phi(\tau)$?
- For a given ("admissible") $g(\tau)$ is there a unique $\Phi(\tau)$?

In other words : Do the second-order statistics of a Hawkes process fully determine the Hawkes process ?

In other words : Do the second-order statistics of a Hawkes process fully determine the Hawkes process ?

This is equivalent to asking whether the equation in Φ

$$g(\tau) = \Phi(\tau) + \Phi \star g(\tau), \quad \text{for} \quad \tau > 0$$

could have more than one positive causal solutions ?

The answer is .... No ! There can be at most one solution !

Thus **The second-order statistics of a Hawkes process do fully determine a Hawkes process**

Proof 1/2

Let Φ_1 and Φ_2 two positive causal solutions of

$$g(\tau) = \Phi(\tau) + \Phi \star g(\tau), \quad \text{for} \quad \tau > 0$$

We set $\Delta(t) = \Phi_1(t) - \Phi_2(t)$, and $B(\tau) = \Delta(\tau) + \Delta(\tau) \star g(\tau)$.
Then $B(\tau)$ is anticausal (i.e., it is 0 for $\tau > 0$)

We get $\hat{B}(z) = \hat{\Delta}(z)(\mathbb{I} + \hat{g}(z))$.

Let us recall that $\hat{g}(z) = \hat{C}^T(z)\Sigma^{-1} - \mathbb{I}$ and

$$\hat{C}(z) = (\mathbb{I} + \hat{\Psi}(-z))\Sigma(\mathbb{I} + \hat{\Psi}^T(z))$$

Thus

$$\hat{g}(z) = (\mathbb{I} + \hat{\Psi}(z))\Sigma(\mathbb{I} + \hat{\Psi}^T(-z))\Sigma^{-1} - \mathbb{I}$$

Thus

$$\hat{B}(z) = \hat{\Delta}(z)(\mathbb{I} + \hat{\Psi}(z))\Sigma(\mathbb{I} + \hat{\Psi}^T(-z))\Sigma^{-1}$$

Proof 2/2

From

$$\hat{B}(z) = \hat{\Delta}(z)(\mathbb{I} + \hat{\Psi}(z))\Sigma(\mathbb{I} + \hat{\Psi}^T(-z))\Sigma^{-1}$$

we get

$$\hat{B}(z)\Sigma(\mathbb{I} - \hat{\Phi}^T(-z)) = \hat{\Delta}(z)(\mathbb{I} + \hat{\Psi}(z))\Sigma$$

In this latter equation

- The left term corrresponds to an anticausal function ($\hat{f}(z) \to 0$, for $|z| \to +\infty$ with $Re(z) > 0$)
- The right term corrresponds to an causal function s($\hat{f}(z) \to 0$, for $|z| \to +\infty$ with $Re(z) < 0$)

From Liouville Theorem we get $\hat{\Delta}(z) = 0$ and consequently

$$\Phi_1(\tau) = \Phi_2(\tau)$$